

UNIVERSIDAD NACIONAL DE PIURA
FACULTAD DE CIENCIAS
ESCUELA PROFESIONAL DE ESTADÍSTICA



**“ APLICACIÓN DE LA REGRESIÓN LOGÍSTICA BINARIA PARA
DETERMINAR LOS FACTORES SOCIALES, FAMILIARES,
ECONÓMICOS Y PSICOLÓGICOS QUE INFLUYEN EN EL
RENDIMIENTO ACADÉMICO DE LOS ESTUDIANTES INGRESANTES
A LA UNIVERSIDAD NACIONAL DE PIURA EN LA FACULTAD DE
CIENCIAS DURANTE SU PRIMER AÑO DE ESTUDIOS 2013 ”**

**TESIS PRESENTADA PARA OBTENER EL TÍTULO DE LICENCIADO
EN ESTADÍSTICA**

PRESENTADO POR: BR. GUSTAVO ADOLFO OLAYA ALBAN.

BR. MARÍA SUSANA COVEÑAS ECHEANDIA.

ASESORADA POR: DR. CARLOS EDUARDO CABRERA PRIETO.

PIURA – PERÚ
2015

7619
OLA



UNIVERSIDAD NACIONAL DE PIURA
FACULTAD DE CIENCIAS
ESCUELA PROFESIONAL DE ESTADÍSTICA

TESIS

“ APLICACIÓN DE LA REGRESIÓN LOGÍSTICA BINARIA PARA DETERMINAR LOS FACTORES SOCIALES, FAMILIARES, ECONÓMICOS Y PSICOLÓGICOS QUE INFLUYEN EN EL RENDIMIENTO ACADÉMICO DE LOS ESTUDIANTES INGRESANTES A LA UNIVERSIDAD NACIONAL DE PIURA EN LA FACULTAD DE CIENCIAS DURANTE SU PRIMER AÑO DE ESTUDIOS 2013”.

Dr. CARLOS EDUARDO CABRERA PRIETO.

Asesor

Br. MARIA SUSANA
COVEÑAS ECHEANDIA.

Br. GUSTAVO ADOLFO
OLAYA ALBAN.



El Jurado Calificador ha dictaminado:

APROBAR POR UNANIMIDAD

La tesis titulada:

" APLICACIÓN DE LA REGRESIÓN LOGÍSTICA
BINARIA PARA DETERMINAR LOS FACTORES
SOCIALES, FAMILIARES, ECONÓMICOS Y
PSICOLÓGICOS QUE INFLUYEN EN EL
RENDIMIENTO ACADÉMICO DE LOS
ESTUDIANTES INGRESANTES A LA UNIVERSIDAD
NACIONAL DE PIURA EN LA FACULTAD DE
CIENCIAS DURANTE SU PRIMER AÑO DE
ESTUDIOS 2013".

Dr. CONRADO SIGISFREDO VARGAS LYNCH.

PRESIDENTE DE JURADO DE TESIS

Dr. GERMAN SANCHEZ MEDINA.

SECRETARIO DE JURADO DE TESIS

M.Sc. LEMIN ABANTO CERNA.

VOCAL DE JURADO DE TESIS

RESUMEN

Toda universidad tiene como objetivo el éxito académico de sus estudiantes, ya que desempeñan un papel protagónico en la sociedad donde están insertadas. Orientados al liderazgo, la realización profesional, el sustento académico adecuado, y la motivación intelectual que genere satisfacción personal en cuanto a las labores y sacrificio que ha significado su responsabilidad durante el periodo de duración de sus estudios universitarios y esto con la finalidad de corrección en las deficiencias del ambiente universitario que se pueden manifestar bajo las formas de rendimiento académico bajo a nivel general de todas las especialidades que conforman este tipo de instituciones y el abandono de la universidad por el estudiante.

Por ello hemos considerado necesario realizar este tipo de investigaciones teniendo como objetivo Determinar los factores sociales, económicos y psicológicos que influyen en el rendimiento académico de los estudiantes ingresantes a la Universidad Nacional de Piura en la facultad de ciencias durante su primer año de estudios 2013, al utilizar la regresión logística binaria. La información se recopiló utilizando una encuesta aplicada a los alumnos de la promoción 2013 de las 5 especialidades de la facultad de ciencias en la Universidad Nacional de Piura; teniendo como resultado los factores y variables que estadísticamente resultan significativas para el cálculo de la Probabilidad (probabilidad de que el estudiante ingresante a la facultad de ciencia de la Universidad Nacional de Piura en el año 2013;salga desaprobado) mediante el modelo de regresión logística binaria son:

- a) Nombre de la especialidad.
- b) Situación económica familiar según el alumno.
- c) Cantidad de presupuesto mensual del alumno.
- d) Opinión familiar sobre su carrera.

De acuerdo a los resultados, el análisis con regresión logística en SPSS mediante el método “Introducir” con el criterio de se ha obtenido el siguiente modelo:

Ecuación:

$$P = \frac{1}{1 + e^{-10.1172 - 0.749X_5 - 1.646X_{10} - 1.004X_{11} + 1.009X_{29}}}$$

Dónde:

X_5 : *Nombre de la especialidad.*

X_{10} : *Situación económica familiar según el alumno.*

X_{11} : *Cantidad de presupuesto mensual del alumno*

X_{29} : *Opinión familiar sobre su carrera*

Los coeficientes β_i son estadísticamente significativos al nivel del 5%. El modelo se ajusta correctamente de acuerdo a la prueba de Bondad de Ajuste de Hosmer y Lemeshow, realizada anteriormente. Los valores numéricos representan el aumento o disminución de la probabilidad individual calculada para cada elemento evaluado, y su expresión numérica esta dado en la ecuación matemática del modelo.

La precisión predictiva global es de 81.0% en el modelo. También observamos que el número de alumnos que salen aprobados en su primer año de estudios son pronosticados correctamente en un 86.4% a este resultado se le conoce como especificidad o susceptibilidad por ser la categoría más común; y sobre el número de alumnos que salen desaprobados en su primer año de estudios son pronosticados correctamente en un 68.8% a este otro valor se le conoce como la sensibilidad porque está dado sobre la categoría de interés o categoría sensible de la variable dependiente.

Es por ello que se recomienda que el procedimiento estadístico en todos sus pasos que lo conforman muestran un efectivo desarrollo con resultados en los indicadores de contraste eficaces, pero debemos considerar las siguientes situaciones, el tamaño de la muestra debería ser más grande se podría extender para más especialidades de la universidad el mismo tipo de diseño para lograr aquel requisito indispensable en la regresión logística binaria, dado que todas las pruebas de contraste se hacen bajo esa condición. La prueba de bondad de Ajuste y la prueba de coeficientes del modelo, así como también la tabla de

clasificación de porcentajes correctos necesitan siempre una significativa disposición en cantidad de elementos para ambas categorías de la variable dependiente, y así entonces estas pruebas se desarrollen correctamente y lograr interpretaciones finales que no violen los supuestos básicos.

PALABRAS CLAVES: El rendimiento académico, inteligencia, edad, regresión logística, curva operativa de rendimiento.

ABSTRACT:

The approach of people to the construction of a knowledge society, understood as the place where citizens take their personal and scientific knowledge based on professional decisions, is mediated by the possibility of access to higher education, since in society it creates an invaluable resource to intensify efforts to build and strengthen social structures that directly benefit the quality of life of people.

Therefore, higher education is necessary for young people to acquire the skills and competencies required to perform adequately in the knowledge society and, at the same time allow them to create bonds of solidarity, ie contribute to social capital nation.

Therefore we considered necessary to perform this type of research using regression models that can be supported on a statistical method to determine those factors that influence student performance so we can correct deficiencies in the universities of our country, whose main objective was determine the social, economic and psychological factors that influence the academic performance of freshmen students of the National University of Piura in the faculty of science during his freshman year 2013, using binary logistic regression. The information was collected using a survey of students of the promotion 2013 of the 5 specialties of the faculty of science at the National University of Piura. The research results show the methodology and procedures used and the variables resulting from the process of investigation; obtaining a model with a good statistical fit. Model variables, the results and conclusions are shown in the development work.

KEY WORDS: Academic achievement, intelligence, age, logistic regression, operational performance curve.

INDICE

| | |
|---|----|
| I. GENERALIDADES..... | 1 |
| 1.1. INTRODUCCIÓN..... | 1 |
| 1.2. EL PROBLEMA DE INVESTIGACIÓN..... | 6 |
| 1.2.1. DESCRIPCIÓN DEL PROBLEMA..... | 6 |
| 1.2.2. FORMULACIÓN DEL PROBLEMA..... | 10 |
| 1.3. JUSTIFICACIÓN..... | 11 |
| 1.4. OBJETIVOS..... | 13 |
| 1.4.1. OBJETIVO GENERAL..... | 13 |
| 1.4.2. OBJETIVOS ESPECÍFICOS..... | 13 |
| 1.5. FORMULACIÓN DE LA HIPÓTESIS:..... | 14 |
| 1.6. LIMITACIONES | 14 |
| 1.7. DELIMITACIONES..... | 15 |
| II. MARCO TEÓRICO..... | 15 |
| 2.1. ANTECEDENTES BIBLIOGRÁFICOS..... | 15 |
| 2.2. EL RENDIMIENTO ACADÉMICO..... | 20 |
| 2.2.1. DETERMINANTES PERSONALES..... | 22 |
| 2.2.2. DETERMINANTES SOCIALES..... | 24 |
| 2.2.3. DETERMINANTES INSTITUCIONALES..... | 25 |
| 2.3. LAS VARIABLES QUE INFLUYEN SOBRE EL RENDIMIENTO ACADÉMICO DEL ALUMNO..... | 27 |
| 2.4. EL RENDIMIENTO EN LAS EVALUACIONES DE APTITUD Y LOS EXÁMENES DE ADMISIÓN..... | 30 |
| 2.5. VARIABLES DE IDENTIFICACIÓN..... | 31 |
| 2.6. VARIABLES PEDAGÓGICAS..... | 36 |
| 2.7. IMPORTANCIA DEL RENDIMIENTO ACADÉMICO..... | 37 |
| 2.8. INTRODUCCIÓN DEL MODELO LOGÍSTICO BINARIO..... | 39 |
| 2.8.1. LA ECUACIÓN LOGÍSTICA..... | 41 |
| 2.8.2. ELEMENTOS DEL ANÁLISIS DE REGRESIÓN LOGÍSTICO..... | 43 |
| 2.8.3. SUPUESTOS DE LA REGRESIÓN LOGÍSTICA..... | 44 |

| | |
|---|----|
| LOS ESTADÍSTICOS ODDS RATIO..... | 45 |
| 2.8.4. SIGNIFICATIVIDAD ESTADÍSTICA DE LOS PARÁMETROS ESTIMADOS..... | 46 |
| 2.8.5. MEDIDAS DE BONDAD DE AJUSTE DEL MODELO..... | 47 |
| ESTADÍSTICO DE WALD..... | 49 |
| PRUEBA DE HOSMER – LEMESHOW..... | 51 |
| PRECISIÓN EN LA PREDICCIÓN..... | 57 |
| REALIZANDO PREDICCIONES..... | 58 |
| CURVA OPERATIVA DE RENDIMIENTO..... | 58 |
| III METODOLOGÍA..... | 59 |
| 3.1. TIPO DE INVESTIGACIÓN..... | 59 |
| 3.1.1. SEGÚN LA NATURALEZA DEL OBJETO DE ESTUDIO..... | 59 |
| 3.1.2. SEGÚN EL MÉTODO DE ESTUDIO DE LAS VARIABLES..... | 59 |
| 3.1.3. SEGÚN EL NÚMERO DE VARIABLES..... | 59 |
| 3.1.4. SEGÚN EL TIPO DE DATOS QUE PRODUCEN..... | 60 |
| 3.1.5. SEGÚN EL TIEMPO DE APLICACIÓN DE LA VARIABLE..... | 60 |
| 3.2. POBLACIÓN DE ESTUDIO..... | 60 |
| 3.3. SELECCIÓN DE MUESTRA..... | 61 |
| 3.4. MÉTODO Y PROCEDIMIENTO DE LA INFORMACIÓN. | 65 |
| 3.5. DEFINICIÓN Y OPERACIONALIZACIÓN DE VARIABLES..... | 67 |
| 3.5.1. VARIABLE DEPENDIENTE..... | 67 |
| 3.5.2. VARIABLE INDEPENDIENTE..... | 68 |
| OPERACIONALIZACIÓN DE LAS VARIABLES DEPENDIENTES E INDEPENDIENTES DE LA ENCUESTA..... | 70 |
| 3.6. CONSTRUCCIÓN DE LA BASE DE DATOS..... | 79 |
| 3.7. PROCEDIMIENTO DE ESTIMACIÓN..... | 80 |
| 3.8. ANÁLISIS DE DATOS..... | 82 |
| IV. ANÁLISIS DE LOS RESULTADOS..... | 83 |
| 4.1. CONSTRUCCIÓN DEL MODELO LOGIT..... | 83 |
| 4.1.1. ESPECIFICACIÓN DEL MODELO DE REGRESIÓN LOGÍSTICA..... | 83 |
| 4.1.2. AJUSTE DEL MODELO DE REGRESIÓN LOGÍSTICA..... | 84 |
| 4.2. MODELO FINAL..... | 91 |

| | |
|--|-----|
| 4.3. ESTIMACIÓN DEL PORCENTAJE DE PERSONAS ENCUESTADAS CLASIFICADAS CORRECTAMENTE CON EL MODELO LOGIT..... | 92 |
| 4.4. ESTIMACIÓN DEL VALOR DE CORTE ÓPTIMO: CURVA COR (CURVA OPERATIVA DE RENDIMIENTO)..... | 93 |
| 4.5. DISCUSIÓN FINAL..... | 99 |
| V. CONCLUSIONES Y RECOMENDACIONES..... | 101 |
| 5.1. CONCLUSIONES..... | 101 |
| 5.2. RECOMENDACIONES..... | 103 |
| VI. BIBLIOGRAFÍA..... | 105 |
| VII. ANEXO DE LA INVESTIGACIÓN..... | 108 |

I. GENERALIDADES

1.1. INTRODUCCIÓN

Toda universidad tiene como objetivo el éxito académico de sus estudiantes, ya que desempeñan un papel protagónico en la sociedad donde están insertadas. Orientados al liderazgo, la realización profesional, el sustento académico adecuado, y la motivación intelectual que genere satisfacción personal en cuanto a las labores y sacrificio que ha significado su responsabilidad durante el periodo de duración de sus estudios universitarios y esto con la finalidad de corrección en las deficiencias del ambiente universitario que se pueden manifestar bajo las formas de rendimiento académico bajo a nivel general de todas las especialidades que conforman este tipo de instituciones y el abandono de la universidad por el estudiante.

Los primeros años de estudios universitarios son de gran dificultad para todos los alumnos. Ya que tendrán que afrontar una realidad distinta a la secundaria, más aun para aquellos alumnos que viven en las provincias de Piura, o ya sea de otro departamento y han venido a Postular a la Universidad Nacional de Piura y debido a la distancia y tiempo tendrán que alquilar un cuarto para poder estar cerca de la universidad. En algunos casos se verán obligados a trabajar y estudiar para poder solventar los gastos; sin embargo al hacer ambas cosas no dedicara el tiempo suficiente a sus estudios. Y al final del ciclo este sacara un ponderado bajo, producto de desaprobado los cursos o retirarse de estos.

Ante la situación descrita como un ejemplo en forma general se hace necesario definir las variables que puedan estar involucradas en este mal rendimiento estudiantil durante el primer año de estudios de su carrera y mediante aquel modelo relacionar estas variables con el rendimiento académico valorado a través de su promedio anual obtenido en el momento específico de nuestra evaluación, que corresponde al primer año de estudio universitarios, de tal forma que al determinarlo se busca

encontrar un modelo estadístico, en este caso un modelo Logit con una alta precisión predictiva, para poder tomar las medidas preventivas, en base a los resultados que se obtuvieron.

El presente trabajo obtuvo como finalidad esclarecer si los factores que determinan el rendimiento académico de los estudiantes universitarios es el grado de satisfacción percibida por ellos mismos y que puede ser identificada mientras se verifica las notas conseguidas en el año de evaluación, o diversas variables como lo son: el sexo, el semestre el que se encuentra cursando y el programa académico en el que están inscritos. Si se alinea el contexto general de este tipo de investigaciones con la preparación de esta tesis, es necesario describir que la causa inmediata de llevar a cabo su aplicación corresponde al querer saber cómo afecta los diferentes conceptos sociales, familiares, económicos, en los estudiantes de la UNP. Se dice que el rendimiento académico no es el producto de una única capacidad, sino el resultado sintético de una serie de factores que actúan en, y desde, la persona que aprende.

Puede afirmarse, en términos educativos, que el rendimiento académico es un resultado del aprendizaje suscitado por la actividad educativa del profesor y producido en el alumno, aunque es claro que no todo aprendizaje es producto de la acción docente. Se le expresa en una calificación cuantitativa y cualitativa, una nota que si es consistente y válida será el reflejo de un determinado aprendizaje y del logro de unos objetivos preestablecidos (Pita y Corengia, 2005). El rendimiento académico, por ser multicausal, envuelve una capacidad explicativa de los distintos factores y espacios temporales que intervienen en el proceso de aprendizaje. Existen diferentes aspectos que se asocian al rendimiento académico, entre los que intervienen componentes tanto internos como externos al individuo. Pueden ser de orden social, cognitivo y emocional, que se clasifican en tres categorías: determinantes personales, determinantes sociales y determinantes institucionales, que presentan subcategorías o indicadores (Garbanzo, 2007). Como egresados de la universidad hemos observado, en los últimos años, que los estudiantes de

la facultad de ciencias, en sus diferentes escuelas presentan dificultades en el análisis y crítica de los temas de discusión, limitaciones en sus conocimientos de ciencias básicas, fallas en la resolución de problemas, deficiente comprensión de lectura, malos hábitos de estudio, deficiente capacidad de síntesis, entre otros, constituyen factores asociados al rendimiento académico bajo que presentan. Es por ello que decidimos realizar el presente trabajo, con el objetivo de identificar aquellos factores sociales, físicos, mentales, económicos y de hábitos de estudio que influyen en el rendimiento académico bajo, con la finalidad de desarrollar estrategias de solución a este problema.

Para la mayoría de personas el éxito no es un concepto claro, muchos no piensan como conquistarlo, quizá piensen que es exclusividad de una minoría, se olvidan que es un derecho que les pertenece a todos, los que luchan por sus metas e ideales, el éxito profesional comienza por un sueño, en donde los deseos debes hacerlos realidad, porque el mayor éxito en la vida será la satisfacción de lo que se haya logrado, siendo feliz con lo que te gusta, el éxito es algo empírico y contrastable, el éxito se hace al andar, en el camino, mientras se avanza, no cuando se llega. Puede ser planeado o suceder espontáneamente, dependiendo de la situación donde ocurra, por lo general, la noción de éxito se relaciona con el ámbito laboral y social, el éxito es lo expuesto al fracaso o la derrota y es, por tanto uno de los sentimientos o fenómenos más positivos que puede experimentar el ser humano, es por ello que la investigación planeada esta propensa a propagar todas estas definiciones para los jóvenes evaluados y poder inferir que características se pueden rescatar de acuerdo, a los conceptos evaluados, en un lineamiento claro del propósito, causal que se quiere analizar.

Todo proceso educativo busca mejorar el rendimiento del estudiante, de ahí su importancia y la necesidad de considerar los factores que intervienen en él. Estos factores, también llamados determinantes del rendimiento académico, son difíciles de identificar, sin embargo,

requieren acotarse para establecer la influencia e importancia que cada uno tiene en el proceso educativo (Tejedor, 2003). El rendimiento académico es un término multidimensional determinado por los diversos objetivos y logros pretendidos por la acción educativa. Desde la perspectiva operativa del término, se define como la “nota o calificación media obtenida durante el periodo universitario que cada alumno haya cursado” (Tejedor, 1998). Las variables y los indicadores que inciden en él se clasifican de diversas maneras, a saber: variables demográficas o de identificación (sexo, edad, estado civil, experiencia laboral), variables académicas (tipos de estudios cursados, curso, opción en que se estudia una carrera, rendimiento previo), variables socio-familiares (estudios de los padres, situación laboral de los mismos, lugar de residencia familiar, lugar de estudio).

Asimismo existen factores inherentes a las condiciones socioeconómicas, familiares e individuales, que hacen que el nivel, de éxito académico por parte de un estudiante universitario sea un proceso multicausal, donde interactúan factores de riesgo (entendidos como las características, del individuo, grupo o comunidad social donde se realizan las actividades educativas universitarias, esta señalan una alta probabilidad de tener un fracaso) y factores protectores institucionales, por sobre todo la cultura educativa, y desempeño de los estudiantes en cuanto su desenvolvimiento, y su desarrollo humano, que pueden mostrar, los verdaderos efectos de este riesgo involucrado.

Entre los factores asociados a las condiciones socioeconómicas se pueden citar las expectativas sociales, y el capital social en cuanto a las condiciones materiales, el capital humano en una valoración de la aptitudes, y condiciones intelectuales, en paralelo con las correspondientes satisfacciones familiares, condiciones de vida, las expectativas, y el factor individual aparecen las percepciones de los jóvenes en cuanto a su desarrollo académico, a las experiencias de su desenvolvimiento universitario, y sus capacidades que deben ser

preparadas y confirmadas hacia sus inclinaciones profesionales y existenciales.

Algunas investigaciones realizadas en países desarrollados en los estudiantes con éxito en el rendimiento académico señalan que el sexo, el rendimiento escolar previo a la universidad y el factor económico son factores de pronóstico importante para el éxito académico: a ellos se agregan otros factores, como la integración social y los hábitos de estudio continuo en casa, es de vital importancia entonces que todo trabajo que se desea realizar de la misma manera considere como factores principales de análisis los mencionados, además de agregar un número de preguntas que representen todos los factores restantes involucrados con las características de la investigación, que según la percepción del estudio y el conocimiento empírico de donde nace este trabajo son causales de existencia de alumnos con rendimiento académico bajo y por ende reflejado en calificaciones malas.

El rendimiento académico es la resultante del complejo mundo que rodea al estudiante, determinado por una serie de aspectos cotidianos : Esfuerzo, capacidad de trabajo, intensidad de estudio, competencias, aptitud, personalidad, atención, motivación, memoria, medio relacional, que afectan directamente el desempeño académico de los individuos. Otros factores adicionales que influyen pueden ser psicológicos o emocionales como ansiedad o depresión manifestados como nerviosismo, falta o exceso de sueño, incapacidad para concentrarse, apatía y en casos extremos, depresión profunda. Existen otros factores como lo económico, social, la cultura o la práctica de deporte que pueden influir en éste.

Asimismo existen factores inherentes a las condiciones socioeconómicas, familiares e individuales, que hacen que el nivel, de éxito académico por parte de un estudiante universitario sea un proceso multicausal, donde interactúan factores de riesgo (entendidos como las características, del individuo, grupo o comunidad social donde se realizan las actividades

educativas universitarias, esta señalan una alta probabilidad de tener un fracaso) y factores protectores institucionales, por sobre todo la cultura educativa, y desempeño de los estudiantes en cuanto su desenvolvimiento, y su desarrollo humano, que pueden mostrar, los verdaderos efectos de este riesgo involucrado.

Entre los factores asociados a las condiciones socioeconómicas se pueden citar las expectativas sociales, y el capital social en cuanto a las condiciones materiales, el capital humano en una valoración de la aptitudes, y condiciones intelectuales, en paralelo con las correspondientes satisfacciones familiares, condiciones de vida, las expectativas, y el factor individual aparecen las percepciones de los jóvenes en cuanto a su desarrollo académico, a las experiencias de su desenvolvimiento universitario, y sus capacidades que deben ser preparadas y confirmadas hacia sus inclinaciones profesionales y existenciales, por todo lo descrito que resume individualmente los factores posibles que guardan relación con el evento de interés en la variable de respuesta, basado en el conocimiento de resultados logrados por investigaciones similares es necesario realizar un procedimiento conjunto que genere indicadores en cuya capacidad conceptual y de argumentación expliquen cual es la relación que existe entre estos factores con la categoría de interés en investigación y proporcionar conclusiones respecto a esos valores numéricos, todo esto en un análisis conjunto de una técnica multivariada como lo es el análisis de regresión logística.

1.2. EL PROBLEMA DE INVESTIGACIÓN

1.2.1. Descripción del Problema

Antes de entrar en el tema, es necesario hacer algunas precisiones. En primer lugar, debe entenderse que cualquier análisis de la situación de

las universidades en nuestro país tiene que hacerse considerando la heterogeneidad que oculta el concepto de la universidad peruana. Es imposible hacer valoraciones generales, pues la situación de unas universidades respecto a otras es del todo dispar e incluso, suelen existir atribuidas diferencias entre facultades de una misma universidad.

Las universidades desempeñan un papel protagónico en la sociedad donde están insertadas. Orientadas al liderazgo, al desarrollo del conocimiento, son escenario de formación técnica y profesional de los miembros de la comunidad a través de la educación, la investigación, la difusión del saber, orientando y apoyando los cambios en diversos espacios nacionales e internacionales. Una visión nueva de la educación debe ser capaz de hacer realidad las posibilidades intelectuales, espirituales, afectivas, éticas y estéticas, que garanticen el progreso de la condición humana, que promueva un nuevo tipo de ser humano capaz de ejercer el derecho al desarrollo justo, equitativo, que interactúe en convivencia con sus semejantes, con el mundo, y que participe activamente de la preservación de sus recursos.

Bajo este contexto, la misión de las instituciones educativas, específicamente de la Universidad, aparte de ser un escenario de práctica e ilustración, también se constituye en un escenario de interacción y formación social para muchos estudiantes. Además de satisfacer y cubrir el reto de formar profesionales, ciudadanos cultos capaces de configurar sociedades solidarias y de progreso, las universidades tienen un nuevo reto que afrontar: Ser un contexto de vida que propicie comportamientos saludables, que influyan en la calidad de vida, no sólo del colectivo de los que viven y trabajan en la universidad sino de toda la sociedad en general.

En línea con los cambios mundiales, el número de universidades peruanas ha crecido significativamente. Según Tubino y Abugattás

(2001), hacia 1960, el Perú contaba con diez universidades, de las cuales solo una era privada. En cambio, a fines del año 2006, las estadísticas de la Asamblea Nacional de Rectores (ANR) revelan que existían 91 universidades (35 estatales y 56 privadas). Estas instituciones ofrecían 162 carreras profesionales en pregrado, de las cuales las dos de mayor oferta correspondían al ámbito económico empresarial: Administración y Contabilidad (Asamblea Nacional de Rectores, 2006).

Antonio Mabres en su artículo: *Problemas y perspectivas de las universidades peruanas*, considera que son cinco los problemas más importantes de la universidad peruana; **bajo nivel académico**, incertidumbre económica, escasa investigación de calidad, poquísimas relaciones de colaboración académica entre universidades, e inexistencia casi total de relación con el sector empresarial.

Si bien el número de universidades ha aumentado, nada garantiza que la calidad de la oferta en educación superior haya crecido a la par o, al menos, se haya mantenido. Además, la brecha que existe entre la educación básica (escolar) y la universitaria es un problema evidente en el país. Los jóvenes suelen llegar a los estudios superiores sin los conocimientos, las habilidades y las actitudes necesarias para aprovechar al máximo el proceso de aprendizaje. “Ello genera severas distorsiones en la actividad lectiva, sobre todo durante los primeros ciclos académicos, en los cuales se tiene que reforzar materias y competencias que deberían haber sido desarrolladas en la educación básica. Las repercusiones de estas deficiencias se observan en la universidad a través de indicadores tales como el alto número de repeticiones en los cursos básicos y el número de ciclos que emplean los alumnos para terminar su carrera” (Ministerio de Educación del Perú, 2006, p.48).

Se debe considerar que el déficit en los aprendizajes, que pueden ser atribuidos a la educación escolar, no necesariamente implica que el estudiante carezca de potencial académico, pero dificulta el proceso de enseñanza-aprendizaje y exigiría ajustes en los diseños curricular y didáctico para mejorar el rendimiento y reducir la probabilidad de desaprobar en las asignaturas universitarias. Por ello, de identificarse con precisión las variables o factores que tienen mayor influencia sobre el rendimiento universitario, las instituciones dedicadas a la educación superior, podrán realizar ajustes pertinentes en las estrategias de selección y nivelación de sus estudiantes.

Asimismo, sería viable introducir mejoras en el proceso de enseñanza en función de las características del alumnado. De este modo, se contribuiría a un aprendizaje más significativo y a la mejora del rendimiento. Es así que surge la necesidad de conocer en qué medida las variables que no están directamente relacionadas con el desarrollo pedagógico de la universidad influyen sobre el rendimiento de sus estudiantes; entre ellas podemos mencionar a las académicas (en especial, las asociadas con las características de la educación básica), psicológicas, socio-familiares y de identificación. En ese sentido, lo que se busca con el presente estudio es identificar, específicamente para las instituciones de educación superior, este caso universidades nacionales y como ejemplo de aplicación información particular de la Universidad Nacional de Piura, cuáles son los factores que garantizan que los estudiantes poseen los conocimientos, habilidades y actitudes necesarias para enfrentar adecuadamente sus estudios universitarios y culminarlos satisfactoriamente.

Por otro lado, se pretende extender a partir de los resultados una dirección y forma de como orientar las políticas y las estrategias de las universidades, con miras a cubrir las expectativas de sus estudiantes y hacer más agradable su proceso de aprendizaje. Finalmente, se espera establecer algunas primeras conclusiones

respecto de las limitaciones de la educación superior esto acorde con las respuestas a las preguntas que indagan sobre la percepción del estudiante en cuanto al nivel y satisfacción de educación que recibe en relación con el ambiente, la infraestructura, el desempeño académico de sus docentes, la gestión de sus autoridades, específicamente, en lo referido a su capacidad global de satisfacción como alumno para permitir un tránsito más adecuado en el quehacer de sus actividades estudiantiles y en consecuencia las mejoras de su nivel académico y formación profesional.

En el plano local y coherente, dándole el énfasis específico a la investigación que se está cursando, podemos describir que en la Universidad Nacional de Piura, se puede notar una deserción inmediata en los primeros semestres de desarrollo académico universitario por parte de los jóvenes ingresantes, a las especialidades que conforman la facultad donde se realiza el estudio, entonces al saberlo de forma subjetiva, es aquella la principal iniciativa de querer verificar si todas las condiciones y factores antes mencionados según la distinta literatura estudiada, corresponden un efecto principal de influencia sobre la variable de respuesta como valoración de aquella deserción para nuestro trabajo de investigación.

1.2.2. Formulación del Problema

¿Cuáles son los factores sociales, familiares, económicos y psicológicos que influyen en el rendimiento académico de los estudiantes ingresantes a la Universidad Nacional de Piura en la facultad de ciencias durante su primer año de estudios 2013, con la utilización de la regresión logística binaria?

1.3. JUSTIFICACIÓN

El rendimiento académico estudiantil es un problema, que a través de los tiempos y en todos los niveles de la educación, ha ocupado la atención de quienes de una u otra forma, están comprometidos en la difícil tarea de educar, sobre todo, en el subsistencia de educación superior, por ser este un nivel terminal en la obtención de un título universitario. Así cualquier investigación que arroje luces sobre él es fundamental, ya que permite una mejor comprensión de la situación. Actualmente, a pesar de la preocupación y los estudios realizados en la búsqueda de soluciones, estos han sido insuficientes y las consecuencias continúan reflejándose en forma negativa en el estudiante. Los programas de formación universitaria en cualquier especialidad están orientados a formar integralmente jóvenes poniéndolos en contacto directo con su contexto laboral y recibiendo además una completa formación teórico-académica paralela a su capacitación práctica. Para alcanzar esta ambiciosa meta las instituciones públicas y privadas de la Región Piura que actúan como centros de aplicación de estos programas, deberían ofrecer una serie de beneficios socioeconómicos a los participantes con mejor rendimiento académico con la intención de estimular y mantener “motivado”, su comportamiento frente a su desarrollo dentro del quehacer universitario.

Es en tal sentido el presente trabajo de investigación titulado: *“MODELO DE REGRESION LOGISTICA BINARIA PARA DETERMINAR LOS FACTORES SOCIALES, FAMILIARES, ECONOMICOS Y PSICOLOGICOS QUE INFLUYEN EN EL RENDIMIENTO ACADEMICO DE LOS ESTUDIANTES INGRESANTES A LA UNIVERSIDAD NACIONAL DE PIURA, FACULTAD DE CIENCIAS DURANTE EL PRIMER AÑO DE ESTUDIOS 2013”*. Se constituye como un mecanismo para lograr determinados objetivos del rendimiento académico, ya que los estudiantes universitarios se ven afectados tanto negativa como positivamente por varios factores en el desempeño académico, estos factores están relacionadas a los aspectos sociales,

económicos y educativos; con los cuales se busca conocer la importancia de tener familia, dinero, motivación, entre otros, con ellos conocer que tanto afecta en el desempeño de los estudiantes.

Los resultados que se obtuvieron del presente estudio sobre el análisis de los factores que determinan el rendimiento académico, permitirán identificar y proponer algunas alternativas de solución para mejorar el nivel del rendimiento académico en los estudiantes de la UNP en el primer año. Asimismo el presente trabajo pretende ser un instrumento para las futuras investigaciones, que tengan la posibilidad de ampliar con mayor profundidad el tema del nivel de rendimiento académico y que contribuyan a la solución de este problema.

Es necesario remarcar que la pretensión académica en realizar este trabajo de investigación se fundamenta principalmente en encontrar las causas verdaderas de forma científica que contribuyen en la deserción académica profesional de la facultad de Ciencias, representada o reflejada mediante sus calificaciones que viene a ser el indicador inicial que genera nuestra variable dependiente en el estudio, entonces se trata de descubrir a través de los métodos cuantitativos quienes son los factores más importantes que influyen en el bajo rendimiento de los estudiantes, por sobre todas las situaciones en sus dos primeros ciclos de estudio universitario, y así propagar alguna estrategia que garantice mejores resultados ante este problema académico que de forma subjetiva parece ser irrenunciable en el mediano plazo, vulnerando así cualquier proyecto de imagen institucional para la facultad.

Otra de las motivaciones en desarrollar este trabajo es la aplicación de los conocimientos recibidos en nuestra etapa de estudiantes, y de esa manera dar utilidad a cada uno de los conceptos teóricos implicados en esta técnica estadística de regresión que participa como modelo de predicción, de la cual nos interesa estudiar cada una de las particularidades específicas que pertenecen a su estructura global, como son: los requisitos que debe

cumplir para su utilización, la validación, y significación del modelo, los coeficientes, los valores predichos, las comparaciones de porcentajes, y otros indicadores específicos que generan la respuesta final de un modelo estadístico adecuado, con un buen ajuste y con buenos niveles de aplicación predictiva, obteniendo también una interpretación correcta de fácil comprensión y de sencillo manejo en cuanto a sus descripciones teóricas que lo fundamentan.

1.4. OBJETIVOS

1.4.1. Objetivo General

Determinar los factores sociales, económicos y psicológicos que influyen en el rendimiento académico de los estudiantes ingresantes a la Universidad Nacional de Piura en la facultad de ciencias durante su primer año de estudios 2013, al utilizar la regresión logística binaria.

1.4.2. Objetivos Específicos

1. Identificar individualmente los factores obtenidos significativos para este modelo de regresión logística binaria, como variables psicológicas, de identificación, socio-familiares, y económicas de los estudiantes ingresantes a la universidad Nacional de Piura, Facultad de Ciencias, en el primer año de estudios 2013, que contribuyen estadísticamente en la probabilidad de encontrar alumnos desaprobados.
2. Calcular la probabilidad para el rendimiento académico de los estudiantes ingresantes de forma individual, y obtener los porcentajes de clasificación correcta para ambas categorías descritas en la variable dependiente, y obtener la capacidad predictiva global.

3. Establecer los coeficientes de razón de oportunidad que afectan a cada variable significativa dentro del modelo utilizado.
4. Realizar la medida de comparación del punto corte óptimo a través de la curva COR, encontrado los valores de sensibilidad y especificidad del modelo y establecer si existe un punto que genere mejoras de predicción para modelos similares en el futuro.

1.5. FORMULACIÓN DE LA HIPÓTESIS

Existen factores sociales, familiares, económicos y psicológicos que influyen en el rendimiento académico de los estudiantes ingresantes a la Universidad Nacional de Piura en la facultad de ciencias durante su primer año de estudios 2013.

1.6. LIMITACIONES

Esta investigación presenta las siguientes limitaciones:

- La falta de información consolidada y estadísticas de los alumnos que inicien sus estudios en la Universidad Nacional de Piura, durante su primer año de estudio 2013.
- La UNP no tiene ninguna base que contenga información académica, socio-económica, psicológica y de identificación general de la población a considerar.

- No existe información consolidada de datos referidos a las características y la oferta curricular de los colegios de procedencia de la población que se quiere estudiar.
- No existe estudio acerca de las características de la educación escolar y el rendimiento académico en los colegios de la Región Piura, en sus diversas provincias y jurisdicciones que la corresponden.

1.7. DELIMITACIONES

Teniendo en cuenta los criterios de investigación científica, esta investigación se ha delimitado de la siguiente manera.

- El periodo de información registrada es para alumnos encuestados.
- El grupo de estudiantes a tomar en cuenta son los alumnos ingresantes a la facultad de ciencias en el año 2013.

II. MARCO TEÓRICO

2.1. ANTECEDENTES BIBLIOGRÁFICOS.

El tema del rendimiento académico de los estudiantes de educación superior, ha sido estudiado por varios investigadores a nivel internacional como nacional. A continuación mostramos algunos estudios los cuales guardan relación con el presente trabajo de investigación, según esas evaluaciones ya realizadas han sido parte principal de donde se generó la idea de investigación, y por lo tanto se describe sus conclusiones y expresiones importantes y resaltantes a las cuales los investigadores han llegado a concluir.

Ministerio de Educación. (1997), en sus publicaciones sobre Reforma Educativa indica la percepción de los directores y profesores sobre los factores que inciden en el rendimiento académico escolar en Bolivia. Los Directores concluyen que los factores que favorecen la calidad de la educación son el apoyo familiar, formación del profesor, material didáctico apropiado, currículo acorde a la realidad, infraestructura adecuada y la administración eficaz; mientras los Directores opinan respecto a los aspectos que desfavorecen la calidad de la educación son la situación económica de la familia, falta de materiales educativos, el salario de los profesores, el tiempo que ven televisión, aprendizaje en lengua no materna y las enfermedades de los estudiantes

Finalmente, los directores como los profesores opinan que la situación económica de la familia es el factor que más desfavorece la calidad de la educación; las enfermedades es un tema priorizado por los profesores, quienes están en contacto más directo con los alumnos. Los directores perciben que la falta de material educativo afecta en mayor medida la calidad de la educación. Por último, tanto los directores como los profesores también destacan que el salario de los profesores afecta negativamente la calidad de la educación, en mayor medida que otros factores, como el aprendizaje en lengua no materna.

Pascarella y Terenzini (1980), en su artículo *“Predicting Resuman Persistente and Voluntary Dropout Decisions from a Theorical Model”* busca determinar si una medida multidimensional de integración social y académica, basada en los elementos conceptuales del modelo de Tinto, podría discriminar, significativamente, entre los estudiantes de primer año que persisten, y los que se retiran voluntariamente. Para el estudio, emplearon 1.447 encuestas aplicadas a los estudiantes que ingresaron al primer año, a la Universidad de Siracusa. Las preguntas estaban diseñadas de tal manera que capturaban información sobre las variables propuestas por Tinto. Mediante la utilización de métodos de la estadística multivariante, como el análisis factorial y el análisis discriminante, obtuvieron, como resultado, que el modelo teórico propuesto por Tinto explica, de una manera adecuada, el fenómeno de la deserción.

El trabajo de *Giovagnoli (2002)*, utiliza modelos de duración, cuya particularidad es que permiten calcular la probabilidad de que un estudiante deserte de la institución, dado que se encuentra en un determinado semestre. Entre los resultados más sobresalientes, vale la pena mencionar que, cuando un estudiante se vincula a la universidad inmediatamente después de finalizada la secundaria, disminuye la probabilidad de desertar. No existen diferencias significativas, en el riesgo de desertar, entre quienes iniciaron y abandonaron ya otra carrera, y los principiantes. El fracaso vocacional no se constituye en una variable explicativa relevante, que explique el riesgo de deserción; por el contrario, el sexo masculino y el estado civil soltero incrementan el riesgo de deserción.

Jadue, G. (2003), en su documento “Transformaciones Familiares en Chile: Riesgo Creciente para el Desarrollo Emocional, Psicosocial y la Educación”, indica que la implicancia de la familia en la tarea educativa comprende la participación activa de los padres en los proyectos educativos de la escuela y en su rol como mediadores del aprendizaje, siendo la madre un fuerte predictor del rendimiento académico. Sin embargo, las familias de nivel socioeconómico bajo, deben lidiar con un conjunto de estrés que dificultando su rol en el apoyo temprano del niño en el plano académico como la inestabilidad laboral, recursos materiales y financieros inadecuados, problemas maritales y familiares.

Castañón, E. (2004) en su investigación sobre la deserción “Deserción estudiantil Universitaria: Una Aplicación de Modelos de Duración en la Facultad de Ingenierías de la Universidad de Antioquia-Colombia”, encuentra cuatro conjuntos de factores como determinantes de la deserción estudiantil. El primer factor es el individual, el cual reúne las características demográficas del estudiante. El segundo factor es el académico, el cual está asociado a variables que contemplan la educación y la orientación previa del estudiante, así como el desempeño académico dentro de la institución. El tercer factor es el socioeconómico, que considera los aspectos laborales, tanto del estudiante como de la familia, así como la dependencia económica; además, incluye el crecimiento económico, medido a través del PIB, como indicador de la situación económica del país. El cuarto factor es el institucional, que contempla las variables asociadas con la adaptación del estudiante al ambiente universitario.

Los resultados más sobresalientes, en esta investigación, se pueden resumir de la siguiente manera: los estudiantes jóvenes, los hombres, los solteros o aquellos estudiantes que han sufrido alguna calamidad familiar, tienen un mayor riesgo de desertar. En términos académicos, la falta de orientación profesional, el haber egresado de un colegio privado, el tener un bajo rendimiento académico, el no tener experiencia en la educación superior o el haber desertado de otra universidad, incrementan el riesgo de deserción.

Esquivel & Rojas (2005), en su documento “Motivos de los Estudiantes de Nuevo Ingreso para estudiar un Post Grado en Educación”, indica la evidencia de las tendencias motivacionales en situaciones educativas del aprendizaje y del rendimiento; en ambos casos el elemento central es la motivación en función de las expectativas para alcanzar una meta y el valor que está asignado para alcanzarlo, es decir, las expectativas de logro son determinadas por las percepciones que las personas tienen de los factores que han intervenido en sus éxitos o fracasos y estas percepciones serían atribuciones causales que dirigen las pautas e intenciones de la conducta de los estudiantes.

Vélez VA. y col 2005. El objetivo del estudio fue determinar y evaluarlos factores que estuvieran incidiendo en el desempeño académico de estudiantes de primer semestre de medicina. Los resultados mostraron que la composición y las relaciones familiares fueron de 3 a 5 personas viviendo bajo el mismo techo (82%) y se reconoce como jefe del hogar y se depende económicamente de uno o ambos progenitores en forma repartida. Se encontraron rasgos de violencia en el 16% de las familias. Los estudiantes realizaban muy pocas actividades durante su tiempo libre, de tipo cultural o deportivo pero si recreativas como ir al cine, reunirse con los amigos, leer y utilizar el Internet como pasatiempo. El 12.8% de las personas trabajaba en su tiempo libre. El análisis mostró que el no leer como pasatiempo, la presencia de violencia intrafamiliar, el haber fumado marihuana, el provenir de un colegio mixto, el no haber realizado estudios profesionales, de quien se depende económicamente, las notas de biología, bioquímica y del promedio trimestral fueron los factores que se asocian con fracaso académico.

Por su parte, **Abarca y Sánchez (2005)** en las investigaciones realizadas sobre los factores generales que inciden en el rendimiento académico, han encontrado según lo expresado por los estudiantes un obstáculo clave en el bajo rendimiento académico son sus deficiencias en el perfil cognitivo, además de conocimientos disciplinarios insuficientes. Definen dos dificultades primordiales: no saben estudiar y no saben aprender. En el terreno de las habilidades de aprendizaje se localizan dos tipos: Competencias para pensar y comprender Vs. Repetir y memorizar en las diversas disciplinas, y El conocimiento de técnicas de estudio, que sobre todo apuntan a cómo revisar la bibliografía y comprenderla.

Bermúdez S., Durán M., y col 2006. Estudiaron algunos factores de riesgo que inciden en el rendimiento académico (estrés, alcohol, redes de amistad, depresión y funcionalidad familiar) en los estudiantes de la Facultad de Medicina de la Universidad de Manizales Colombia. Se aplicó una encuesta transversal anónima en donde se establecieron características sociodemográficas, culturales, escolares y motivacionales, así como la presencia de estrés, depresión, disfunción familiar y consumo de sustancias. Entre los estudiantes analizados en cuanto a estos factores demográficos no hubo ninguna diferencia estadísticamente significativa con el rendimiento académico. Se encontró relación significativa entre el rendimiento académico y estrés, con depresión, consumo de alcohol y funcionalidad familiar. Se encontró como significativo que los estudiantes con bajo nivel de estrés tenían un promedio de nota de 3.54, mientras que los que presentaban nivel alto de estrés tenían un promedio de 3.39 ($p=0.012$). Análogamente, el 32.7% de los estudiantes que perdieron una materia el semestre anterior presentaban estrés moderado, proporción significativamente superior a lo que ocurre entre quienes no perdieron materias 12.2%, ($p<0.001$).

Contreras K. y col. 2008. Investigaron sobre los factores asociados al bajo rendimiento académico en estudiantes de psicología en una universidad privada de Barranquilla (Colombia). Se entrevistó a 38 estudiantes que asistían a un programa exclusivo para aquellos con muy bajo rendimiento académico y que estaban en riesgo

de salir del programa o habían sido remitidos por los profesores. Los resultados indican que los estudiantes manifiestan limitaciones en sus competencias de concentración y atención, en la distribución del tiempo, en la deficiente utilización de técnicas de estudio, y en la poca asistencia a clases por la baja motivación en las actividades académicas. La relación con su familia la consideraron satisfactoria. En todos estos factores se observa que el contexto juega un papel importante. Se considera que la familia, las variables socioeconómicas y, sobre todo, la atención que reciben los estudiantes por parte de los docentes de las instituciones (diferencia el mayor apoyo que reciben en las instituciones privadas) son factores fundamentales para entender el éxito o fracaso de los estudiantes.

Lindblad F. y col. 2011. El objetivo de este estudio fue evaluar la influencia de la mala calidad del sueño, el estrés percibido y las actitudes alimentarias negativas para predecir dificultades de concentración en la escuela. La muestra estuvo conformada por 1.124 estudiantes (de sexto a noveno grado). El análisis de regresión logística con ajuste mutuo para todos los predictores, así como para el grado y el sexo dio un Odds ratio (OR) para el componente de estrés de 3.05 (intervalo de confianza del 95% [IC 95] = 2.13 - 4.38); OR = 1.70 (1.20-2.42) para las actitudes alimentarias escolares, y OR = 2.57 (1.78-3.71) para las dificultades para dormir, concluyéndose que el sueño, el estrés y las actitudes alimentarias parecen predecir independientemente los problemas percibidos de falta de concentración en la escuela.

2.2. EL RENDIMIENTO ACADÉMICO

En el rendimiento académico inciden muchas variables. Se sostiene la existencia de conexiones significativas entre rendimiento e inteligencia. El factor psicopedagógico que más peso tiene en la predicción del rendimiento académico es la inteligencia y dentro de ésta el factor verbal es el que incide directamente en el rendimiento académico del alumno en todas las áreas del currículo. Sin embargo, no todo son variables individuales. Se evidencia también que las relaciones de amistad constituyen un apoyo que aumentaría la probabilidad de éxito académico, aunque esto aplica más para los hombres que para las mujeres.

Medir el desempeño académico en la universidad no es una tarea sencilla y a lo largo del tiempo la literatura empírica ha encontrado diversas formas de aproximarlos. El problema surge básicamente por la naturaleza propia del proceso educativo, que transforma a los individuos en personas más calificadas. En este punto, sería deseable computar las diferentes calidades de instrucción recibidas. Sin embargo, una alta proporción de los elementos que la componen no pueden observarse o son de difícil cuantificación. La opción más utilizada es considerar las calificaciones obtenidas en un lapso de tiempo determinado y computar el promedio. La idea subyacente es que los alumnos con mejores calificaciones habrían incorporado mayor capital humano en un período determinado. Si se comparan dos estudiantes similares en otros aspectos excepto por el promedio obtenido en los primeros cuatro años de carrera, podría decirse que ha adquirido mayores conocimientos el individuo con calificaciones superiores. Sin embargo, es importante notar que la forma de evaluación y las notas que se asignan podrían diferir entre cursos debido a la subjetividad de cada profesor y esto no puede ser controlado por el indicador.

De esta forma, son numerosas las causas que pueden justificar y concurrir en el rendimiento académico del estudiante universitario, es decir, el bajo rendimiento puede ser ocasionado por factores académicos o personales. Pero también puede ocurrir al contrario: los factores personales son susceptibles de ser afectados por un bajo rendimiento académico.

El rendimiento académico, por ser multicausal, envuelve una enorme capacidad explicativa de los distintos factores y espacios temporales que intervienen en el proceso de aprendizaje. Diferentes aspectos se asocian al rendimiento académico, entre los que intervienen componentes tanto internos como externos al individuo. Pueden ser de orden social, cognitivo y emocional que se clasifican entre sí en varias categorías:

- Determinantes personales
- Determinantes sociales
- Determinantes institucionales

2.2.1. Determinantes personales

Incluye aquellos factores de índole personal, cuyas interrelaciones se pueden producir en función de variables subjetivas, sociales e institucionales.

- Competencia Cognitiva**, definida como la autoevaluación de la propia capacidad o habilidad intrínseca del individuo para cumplir una determinada tarea cognitiva, su percepción sobre su capacidad y habilidades intelectuales. Esta característica se cimienta en la repetición de experiencias de éxito que el individuo haya alcanzado observando el logro en los demás, críticas verbales positivas y los estados emocionales y somáticos. Cada individuo tiene su propia interpretación de sus estados emocionales y somáticos, la cual repercute en la construcción de la competencia cognitiva. Estudios llevados a cabo con estudiantes universitarios pusieron en evidencia que aquellos estudiantes con creencias de autoeficacia académica positiva se asocian con resultados de éxito académico.
- Motivación**, la cual puede ser intrínseca, extrínseca, atribuciones causales y percepciones de control. La motivación intrínseca es definida como un estado psicológico relacionado con los estudios que es positivo y significativo; se caracteriza por el vigor, dedicación y absorción que presenta el individuo por la labor realizada. La motivación extrínseca, se relaciona con aquellos factores externos al estudiante, cuya interacción con los determinantes personales da como resultado un estado de motivación. Entre estos factores puede estar el tipo de universidad, los servicios que ofrece la institución, el compañerismo, el ambiente académico, la formación del docente, condiciones económicas, entre otras. Se ha demostrado que asumir

que los resultados académicos se deben a la propia capacidad y esfuerzo, influye en el logro de buenos resultados académicos. Las percepciones de control, constituye la percepción del estudiante sobre el grado de control que se ejerce sobre su desempeño académico.

- c. **Las condiciones cognitivas**, son las estrategias de aprendizaje que el estudiante lleva a cabo relacionadas con la selección, organización y elaboración de los diferentes aprendizajes. Se definen como condiciones cognitivas del aprendizaje significativo. El uso de mapas conceptuales, hábitos de estudio, horas asignadas al estudio, las prácticas académicas son algunas estrategias de aprendizaje utilizadas por los estudiantes.
- d. **El autoconcepto académico**, conjunto de percepciones y creencias que una persona posee sobre sí misma. La capacidad percibida por parte del estudiante, el rendimiento académico previo y creer que la inteligencia se desarrolla a partir del esfuerzo académico, contribuyen a mejorar un auto concepto académico positivo.
- e. **La autoeficacia percibida**, la ausencia de autoeficacia en los estudiantes está asociada a estados de agotamiento, desinterés y falta de proyección con sus estudios y es conocido como Burnout, que es la fatiga o la sensación de estar “quemado” por las actividades académicas.
- f. **Bienestar psicológico**, se ha demostrado que cuanto mayor rendimiento académico haya habido en el pasado mayor será el bienestar psicológico en el futuro, lo cual incidirá en el rendimiento académico. La satisfacción hace referencia al bienestar del estudiante en relación con sus estudios, e implica una actitud positiva hacia la universidad y la carrera.

- g. **Asistencia a clases**, es una de las variables que según diversos autores es una de las más significativas. Cuanta mayor asistencia mayor será el rendimiento académico.
- h. **Inteligencia**, es un buen predictor de los resultados académicos. Incluye comprensión verbal, razonamiento matemático.
- i. **Nota de acceso a la universidad**, aquellas universidades donde se practica un proceso de selección pormenorizado y completo en cuanto a los requisitos de entrada, presentan un bajo grado de fracaso ente sus estudiantes.

2.2.2. Determinantes sociales

Constituyen aquellos factores asociados al rendimiento académico de índole social que interactúan con la vida académica del estudiante. Entre ellas tenemos a las desigualdades sociales, el entorno familiar, el nivel educativo de los progenitores o adultos responsables del estudiante, nivel educativo de la madre, contexto socioeconómico y variables demográficas.

- a. **Desigualdades sociales**, factores como la pobreza y la falta de apoyo social está relacionado con el fracaso académico, sin embargo no hay una correspondencia estricta.
- b. **Entorno familiar**, entendido por el conjunto de interacciones propias de la convivencia familiar, que afectan el desarrollo del individuo manifestándose también en la vida académica. Un ambiente familiar propicio marcado por el compromiso así como una convivencia familiar democrática, inciden en un adecuado desempeño académico.
- c. **Nivel educativo de los progenitores**, se plantea que la estimulación educativa de parte de los padres con mayor nivel sociocultural y

educativo es la responsable de las diferencias de los resultados académicos en estudiantes de diferentes niveles económicos.

- d. **Nivel educativo de la madre**, cuanto mayor sea el nivel académico de la madre, mayor percepción de apoyo hacia sus estudios tienen los hijos e hijas, lo cual se refleja en el rendimiento académico. Asimismo cuando mayor nivel educativo tiene la madre, mayores exigencias académicas les plantea a sus hijos, sustentado en la mayor posibilidad de éxito futuro.
- e. **Contexto socioeconómico y variables demográficas**, varios estudios indican una relación con el rendimiento académico, pero sin poder atribuirle una causalidad directa y estricta.

2.2.3. Determinantes institucionales

Estos factores tienen gran importancia, debido a que desde el punto de vista de tomas de decisiones pueden en cierta manera ser establecidos, controlados o modificados. Entre ellos se encuentran: elección de estudios, condiciones institucionales, complejidad de los estudios, ambiente estudiantil, la presencia de servicios institucionales de apoyo, relación estudiante profesor entre otros.

- a. **Elección de los estudios**, la posibilidad de ubicarse en una carrera universitaria escogida en primera opción de interés del estudiante, repercute en el rendimiento académico del estudiante.
- b. **Condiciones institucionales**, muchos investigadores han demostrado que elementos como: condiciones de las aulas, servicios, plan de estudios y formación de los profesores, pueden ser tanto obstaculizadores como facilitadores del rendimiento académico. Estos factores son de gran importancia porque pueden de alguna manera modificarse y controlarse, por ejemplo el mejoramiento de aulas, el horario de los cursos, el tamaño de alumnos por asignaturas, etc.

- c. **Servicios instituciones**, hace referencia a todos aquellos servicios que la institución ofrece al estudiante según su condición económica, por ejemplo: becas, préstamo de libros, asistencia médica, apoyo psicológico, etc.
- d. **Ambiente estudiantil**, un ambiente con excesiva competitividad con los compañeros puede ser un factor tanto obstaculizador como facilitador del rendimiento académico. La solidaridad, el compañerismo y el apoyo social son elementos que inciden positivamente.

Toda la descripción anterior se manifiesta como una percepción generalizada acerca de los factores que se debieran tomar como importantes dado que su explicación demuestra que tiene una relación muy fuerte con los conceptos que se tienen como base inicial de la investigación, en el marco teórico específico de factores considerados de dará una mejor representación ordenada, precisa, detallada y específica de cada una de las preguntas consideradas inmersas lógicamente en los factores analizados.

El presente estudio explicará los resultados académicos, a través de la calificación final (promedio ponderado acumulado) y los créditos acumulados en el primer año de estudios en la Universidad Nacional de Piura, en función de las variables no pedagógicas propuestas por Tejedor (2003): Académicas, psicológicas, socio-familiares y de identificación. Por ello, a continuación, se describirán las teorías educativas que analizan la influencia de estas variables sobre el rendimiento académico en una evaluación, en una asignatura, en un conjunto de asignaturas de una institución o de un sistema educativo. Adicionalmente, se examinará la relación entre las variables pedagógicas y el rendimiento, pues así se

tendrá una idea sobre aquellos factores explicativos que se estarían dejando de lado.

Para sustentar las teorías que se presentará, se acudirá a los resultados de investigaciones fundamentalmente cuantitativas, aunque también se sintetizarán las conclusiones de estudios cualitativos. Se debe señalar que, cuando se trata de identificar y comprender la naturaleza de los factores o variables que son percibidos por estudiantes y/o profesores como potenciales predictores del rendimiento académico, los trabajos cualitativos suelen usar encuestas aplicadas a muestras que no necesariamente poseen representatividad estadística, así como grupos focales y entrevistas a profundidad. En cambio, los estudios cuantitativos acuden a metodologías de tipo experimental y no experimental. A su vez, las investigaciones no experimentales se apoyan en, por lo menos, una de las siguientes tres técnicas de análisis de estadístico: descriptivo, correlacional o regresión.

2.3. LAS VARIABLES ACADÉMICAS QUE INFLUYEN SOBRE EL RENDIMIENTO ACADÉMICO DEL ALUMNO

Las variables académicas

En el Perú, diversas investigaciones regresionales han analizado los resultados escolares. Estos trabajos han identificado que el rendimiento en matemáticas aumenta con el nivel de asistencia a la escuela (Cueto, Jacoby y Pollitt, 1997) y disminuye en los estudiantes repitentes (Caro, 2003). El rendimiento en matemáticas también se incrementa en los centros que tienen biblioteca (Caro, 2003) o mayor número de libros y de computadoras; así como en aquellos que carecen de aulas multigrado (Valdivia y León, 2007).

Asimismo, obtienen un mayor rendimiento los centros educativos de gestión no estatal y aquellos en los cuales el director del colegio posee título universitario (Caro, 2003). También, se han encontrado efectos positivos sobre el rendimiento en matemáticas como resultado de la

presencia de docentes con título universitario en el aula (Caro, 2003, y Valdivia y León, 2007). Igualmente, se ha determinado que los resultados en matemáticas de los escolares peruanos mejoran cuando el estudiante afirma entender el curso y se dedica un mayor número de horas a hacer las tareas de matemáticas (Caro, 2003). Los mismos efectos se presentan cuando el estudiante tiene una actitud positiva hacia las asignaturas de matemáticas (Caro, 2003, y Ministerio de Educación del Perú, 2004).

Sobre el rendimiento escolar en lenguaje, en el Perú, se ha encontrado que influye el nivel de asistencia a la escuela (Cueto, Jacoby y Pollitt, 1997). También afecta de modo positivo el hecho de estudiar en colegios con mejores recursos (profesores titulados, mayor número de computadoras y libros, la ausencia de aulas multigrado); mientras que el aprendizaje del idioma nativo (diferente al castellano) genera un rendimiento más bajo en las pruebas de comunicación (Valdivia y León, 2007). Parte importante de las variables mencionadas extienden su efectos hasta la educación superior, tal es el caso del rendimiento escolar o las características del colegio de procedencia. Adicionalmente, los investigadores suelen considerar un conjunto de variables académicas asociadas netamente a la educación superior, entre las cuales destacan los resultados en los exámenes de admisión o las pruebas de selectividad, el rendimiento previo en la misma universidad, la asistencia a clases, las horas dedicadas al estudio, el apoyo de tutores, el programa o carrera al que pertenece el estudiante, entre otras.

➤ **El tipo de escuela secundaria**

La literatura empírica ha investigado en varias oportunidades el efecto del tipo de escuela secundaria sobre del desempeño académico. El objetivo es evaluar si ciertas características específicas de los establecimientos de educación media como su religión, tipo de gestión (pública o privada), etc. modifican la incorporación de conocimientos de los alumnos en esa instancia, induciendo desempeños diferenciales luego de haber ingresado en la universidad. Birch y Miller (2006) analizan la influencia del tipo de escuela secundaria en Australia sobre toda la distribución condicional de

rendimiento, concluyendo que las escuelas católicas o independientes impactan negativamente sobre el desempeño de los alumnos en todos los cuantiles donde el efecto resulta relevante. Di Gresia, Porto y Ripani (2002) presentan evidencia para Argentina señalando que los alumnos provenientes de escuelas privadas son más exitosos en la universidad que los que asistieron a escuelas públicas. Por último Di Gresia, Fazio, Porto, Ripani y Sosa Escudero (2005) concluyen que la educación media privada es un factor relevante sólo para los alumnos ubicados en la cola inferior de la distribución condicional de rendimiento.

➤ **El desempeño en la escuela secundaria**

Muchas universidades prestigiosas del mundo consideran especialmente las calificaciones obtenidas en la escuela secundaria (o resultado de tests específicos) entre los elementos que evalúan para decidir la admisión de los estudiantes en sus programas. La idea subyacente es que esta medida aproxima, al menos parcialmente, el capital humano acumulado por el estudiante y su capacidad de aprendizaje. Para ciertas instituciones de formación superior distinguir esta característica en los aspirantes resulta crucial, ya que les permite seleccionar aquellos que cuentan con mayor potencial y mantener de esta manera su excelencia y reputación. Por otra parte, al diseñar políticas universitarias también podría ser útil conocer la influencia del promedio en la escuela secundaria sobre el rendimiento en estudios de nivel superior. Ello permitiría identificar los estudiantes en “riesgo” y dirigir acciones orientadas a reducirlo. Diversos trabajos han investigado esta relación encontrando un efecto positivo y significativo en la mayoría de los casos. Betts y Morrel (1999) proveen evidencia de que tanto el promedio en la escuela secundaria como las calificaciones obtenidas en los tests de habilidades que rinden los alumnos antes de ingresar, favorecen su desempeño en la universidad. Birch y Miller (2006) encuentran efectos positivos y significativos en toda la distribución condicional de rendimiento, exceptuando la cola superior donde se ubican los estudiantes que muestran un desempeño muy por encima del promedio.

➤ **El desempeño en el ciclo de formación inicial**

Existe alguna evidencia de que el desempeño en los cursos de formación inicial (que se dictan en los primeros meses de la carrera) predice buena parte del éxito o fracaso académico del alumno en los años posteriores. Como se señalara previamente, los resultados obtenidos en esos meses posiblemente dependan de los conocimientos adquiridos por el estudiante en instancias previas a la universidad y de su potencial de aprendizaje. Sin embargo controlando por el desempeño en la escuela secundaria, es decir considerando alumnos con alcances similares en esa etapa, esta medida ofrece una visión de la capacidad de respuesta del estudiante al nuevo ámbito de estudio sin duda más exigente. Es decir, permite considerar explícitamente dificultades propias del ámbito universitario que exige mayor dedicación, concentración y esfuerzo. Porto y Di Gresia (2004) analizan este efecto en los alumnos de la cohorte 2000 de la Facultad de Ciencias Económicas de la Universidad Nacional de La Plata encontrando una influencia positiva y significativa en diversas instancias de la carrera.

2.4. EL RENDIMIENTO EN LAS EVALUACIONES DE APTITUD Y LOS EXÁMENES DE ADMISIÓN

Existen estudios que operacionalizan el rendimiento previo mediante los resultados de las evaluaciones nacionales que se aplican en diversos países para medir el nivel la aptitud y/o de conocimientos de los escolares cuando terminan la educación secundaria. Los resultados de estas pruebas normalmente son considerados por las universidades de los países en cuestión para seleccionar a sus alumnos. Otras investigaciones relacionan el rendimiento universitario con los resultados de los exámenes de admisión que rinden los postulantes a las universidades o a las instituciones que ofrecen estudios post-secundarios.

2.5. VARIABLES DE IDENTIFICACIÓN

Tejedor (2003) señala que, como variables identificación, se incluyen aquellos determinantes del rendimiento académico que identifican o caracterizan al alumno. Específicamente se refiere a la edad y al sexo del estudiante. Sin embargo, es posible encontrar investigaciones que consideran la raza o del estatus residencial como variables de identificación.

➤ El Sexo

Conviene aclarar que los estudios sobre rendimiento académico suelen considerar al sexo del estudiante como un identificador de su género. Sin embargo, para parte importante de científicos sociales contemporáneos, el género es una construcción social compleja que va más allá de las diferencias biológicas entre ambos sexos. Por ello, si se evalúan las diferencias en rendimiento entre hombre y mujer, en un sentido de identificación, es más pertinente especificar al sexo como la variable de análisis.

Luego de analizar la relación entre sexo y rendimiento escolar, los estudios realizados en los países desarrollados suelen concluir que las mujeres rinden mejor que los hombres. En cambio, en los países latinoamericanos, se ha encontrado que los varones tienen un mejor rendimiento. La desventaja de las mujeres también es recurrente en las investigaciones desarrolladas en el Perú, donde Cortez (2001) y Cueto (2004) concluyeron, respectivamente, que el atraso y la deserción escolar eran más probables en estudiantes de sexo femenino. En el Perú, también se ha encontrado que los hombres tienden a rendir mejor que las mujeres en matemáticas (Caro, 2003, el Ministerio de Educación del Perú, 2004, y Valdivia y León, 2007). En el ámbito de la educación superior, la mayoría de las investigaciones sobre el rendimiento académico apuntan a un mayor éxito entre las mujeres (Tejedor, 2003).

En varios estudios se evidencia la existencia de diferencias significativas en el desempeño académico en la universidad por el género de los estudiantes. Estudios previos han encontrado en varios países que las mujeres se desempeñan mejor en la universidad. Betts y Morrel (1999) en un estudio realizado para 5000 alumnos de la universidad de California encuentran que el rendimiento de las mujeres es superior a los hombres. En el Perú, Torres, Lajo, Campos y Riveros (2007), un estudio para relacionar el rendimiento académico de los alumnos con la percepción de la calidad académica de sus docentes encontraron diferencias significativas en función del sexo, de modo tal que las mujeres tenían calificaciones promedio más elevadas que los varones.

➤ **La edad**

La influencia de la edad en el rendimiento universitario no parece a priori predecible. Podría pensarse que los alumnos más jóvenes tienen menos responsabilidades (económicas, laborales, familiares, etc.) que sus pares de mayor edad. Además posiblemente preserven el ritmo de estudio de la secundaria. Por ello, quizás estén más motivados o focalizados en el estudio y se desempeñen mejor. Sin embargo, uno también podría plantearse que a causa precisamente de esas mayores responsabilidades sean los alumnos más grandes los más motivados, aprovechando al máximo cada minuto invertido en estudiar y por tanto muestren mejor rendimiento. Las investigaciones empíricas no han sido concluyentes. Naylor y Smith (2004) muestran que la edad influencia positivamente el desempeño. Porto y Di Gresia (2001) encuentran evidencia a favor de un efecto negativo. Giovanoli (2002) señala que la edad aumenta el riesgo de desertar pero no influye en la probabilidad de graduarse. Por último Di Gresia, Fazio, Porto, Ripani y Sosa Escudero (2005) encuentran que el coeficiente estimado se mueve monotónicamente de negativo a positivo entre cuantiles lo cual asocian con un fuerte efecto escala sobre el rendimiento.

En la práctica, diversas investigaciones han concluido que la edad, como una medida de madurez, contribuye a un mayor rendimiento. En efecto, Didia y Hasnat (1998) encontraron que la edad impactaba sobre el desempeño en un curso universitario de introducción a las finanzas, aunque de modo marginal. Por su parte, Di Gresia, Porto y Ripani (2002) determinaron que la mayor edad influía positivamente sobre la cantidad de materias aprobadas de los estudiantes de universidades públicas argentinas. A su vez, Hakkinen (2004) estimó que, en las carreras de ingeniería y ciencias sociales, los estudiantes que ingresaban a la universidad entre los 21 y 23 años tenían un mejor rendimiento que quienes comenzaban sus estudios entre los 18 y 20 años.

➤ **El estado civil**

Algunos trabajos han estudiado el efecto del estado civil del alumno sobre su desempeño en la universidad. Una de las hipótesis discutidas es que los alumnos casados serían personas felices y con mayor contención afectiva por lo cual, tal vez posean mayor capacidad de concentración para incorporar nuevos conocimientos. Sin embargo, también podría argumentarse (al igual que en el caso de la edad) que los alumnos solteros tienen menos responsabilidades y disponen de más tiempo para dedicar al estudio. Debe notarse que el estado civil está muy ligado a la edad del estudiante. Porto y Di Gresia (2001), luego de controlar por la edad del estudiante, encuentran que su estado civil tiene un impacto sólo en una de las tres medidas de rendimiento 16 analizadas. En particular sus hallazgos indican que permanecer soltero afecta positivamente el rendimiento. Giovanoli (2002) encuentra que los alumnos solteros tienen mayor riesgo de desertar pero a su vez mayor probabilidad de graduarse.

➤ **El lugar de nacimiento**

La nacionalidad es otro factor que podría inducir diferentes rendimientos. Sin embargo la dirección del efecto no resulta obvia a priori. Por un lado podría pensarse que los alumnos extranjeros tal vez enfrenten problemas de adaptación (especialmente con las costumbres y el lenguaje) o

simplemente sufran el desarraigo y ello dificulte su aprendizaje. Sin embargo también podrían ser alumnos sumamente focalizados y con objetivos más claros. Los estudios no han sido concluyentes. Betts y Morrel (1999) encuentran que la nacionalidad tiene un efecto positivo, esto es, los extranjeros muestran un peor desempeño. Sin embargo, la magnitud del efecto no parece importante y deja de ser significativo cuando se controla por los recursos de la institución y las características socioeconómicas de su entorno. Di Gresia, Porto y Ripani (2002) utilizando información de un conjunto de universidades argentinas concluyen que los extranjeros se desempeñan mejor en 13 casos y peor en 16 casos. No obstante, el coeficiente encontrado es pequeño y resulta no significativo en varias oportunidades. También podría ser relevante investigar si los alumnos nacidos en la ciudad donde se ubica la universidad se han adaptado mejor al entorno y son capaces de incorporar conocimientos con mayor facilidad.

► El tipo y lugar de residencia

La influencia del tipo y lugar de residencia tampoco resulta intuitivamente clara a priori. Por un lado uno podría postular que los alumnos que conviven con sus padres cuentan con mayor apoyo, contención afectiva y tienen menos responsabilidades (económicas y relativas al funcionamiento diario del hogar). Ello les permitiría avocarse sin preocupaciones a sus estudios y aprovechar al máximo su potencial de aprendizaje. En este caso, esperaríamos observar un mejor desempeño en los estudiantes de este grupo. Sin embargo podría suceder que, precisamente por tener a su cargo varias responsabilidades, los estudiantes que viven fuera del hogar paterno hayan alcanzado una madurez superior, tengan bien definidos sus objetivos y muestren una mayor focalización en sus estudios cumpliendo exitosamente los requerimientos académicos exigidos por la universidad. La evidencia empírica parece apoyar, aunque no de manera concluyente, la segunda hipótesis. Naylor y Smith (2004) encuentran un mejor rendimiento en los alumnos que viven en la universidad respecto de los que viven con sus padres. Giovagnoli (2002) muestra que residir con la

familia impacta positivamente sobre el riesgo de deserción y no influye sobre la probabilidad condicional de graduarse. Di Gresia, Porto y Ripani (2002) encuentran que mudarse del hogar paterno durante el período lectivo mejora el rendimiento. Otro factor clave podría ser residir en el mismo lugar donde se ubica la universidad. En este sentido los costos de traslado que incluyen tiempo, desgaste físico y dinero podrían reducir la capacidad de aprendizaje. Naylor y Smith (2004) proveen evidencia a favor de que el efecto de residir lejos del establecimiento educativo, al que asiste el individuo, es negativo. En línea con estos hallazgos Di Gresia, Porto y Ripani (2002) concluyen que viajar entre jurisdicciones para asistir a la universidad tiene un impacto negativo sobre el desempeño de los alumnos.

➤ Variables socio-económicas

Entre las variables socio-económicas que suelen usar los investigadores destacan los ingresos y el nivel socio-económico de la familia, así como la composición del hogar (Vidal, 1994). Respecto de los niveles socio-económicos (NSE), Fernández y Perera (2004) han identificado tres grandes grupos de estrategias para la definición de los grupos socio-económicos que son usados para la estratificación social. El primer grupo de estrategias se basa exclusivamente en la ocupación o profesión de las personas. El segundo grupo define a las categorías o “clases” sociales de acuerdo con la percepción subjetiva de su status. El tercer grupo está conformado por aquellas estrategias que combinan características o atributos de la familia, que a su vez se dividen en dos grupos: las variables relacionadas con la ocupación del jefe del hogar y las relacionadas con el nivel de confort o equipamiento del hogar. Por otro lado, para operacionalizar la composición o estructura del hogar, se acude a la cantidad de miembros de la familia y a su porcentaje de adultos. Otra manera de aproximarse a la estructura del hogar es a través del número de progenitores. En este caso, se suele diferenciar las familias biparentales de las monoparentales. Este último tipo de familia se refiere “a la convivencia de al menos un menor con uno solo de sus progenitores, el

padre o la madre” (Iglesias, 1998, 113). El estudio del impacto del número de progenitores sobre el rendimiento es importante pues, en las investigaciones, se suele encontrar que el bajo logro académico y los problemas de conducta en la escuela se relacionan con la procedencia de hogares monoparentales o de otras estructuras familiares, diferentes a las biparentales. Adicionalmente, los estudios también concluyen que las dificultades en materia de conducta y rendimiento escolar se presentan en hogares donde el alumno reside con su madrastra y/o su padrastro (Zill, 1996).

➤ Variables socio-culturales

El escenario socio-cultural puede entenderse como “un entorno espacio-temporal que contiene un rico entramado de relaciones personales, con actores dotados de intenciones, motivos y metas, que realizan actividades y tareas significativas para la cultura y que, siguiendo determinados formatos interactivos y tipos de discurso, negocian una representación compartida del contenido de las mismas” (Rodrigo y Acuña, 1998, 262). En ese sentido, las variables socio-culturales propuestas por los investigadores que analizan el rendimiento académico procuran evaluar el nivel cultural de los padres y los hermanos del estudiante (Vidal, 1994). Detrás de este intento, está el reconocimiento implícito de que las interacciones entre los distintos miembros de la familia determinan el clima emocional del hogar y, de modo paralelo, ofrecen un modelo social en miniatura (Jiménez, 2006). Entre las diversas variables socio-culturales consideradas por los investigadores educativos, destacan los valores, la religión y el desarrollo lingüístico de la familia. También es común que se analicen las diferencias culturales entre los hogares urbanos y rurales.

2.6. VARIABLES PEDAGÓGICAS

Las modalidades de enseñanza son los distintos escenarios en donde se desarrollan las actividades que realizan el profesorado y el alumnado a lo largo de un curso. Se diferencian entre sí en función de los propósitos de

la acción didáctica, las tareas a realizar y los recursos necesarios para su ejecución. Estos escenarios pueden ser presenciales; es decir, los que incluyen actividades que reclaman la intervención directa de profesores y alumnos como son las clases teóricas, los seminarios, las clases prácticas, las prácticas externas y las tutorías. También pueden ser modalidades no presenciales, cuando se trata de actividades que los alumnos pueden realizar libremente bien de forma individual o mediante trabajo en grupo (De Miguel, 2006). Si bien las variables pedagógicas no serán medidas en el presente estudio, conviene mencionar que la investigación didáctica ha identificado un conjunto de acciones docentes que favorecen un buen rendimiento en el ámbito universitario.

2.7. IMPORTANCIA DEL RENDIMIENTO ACADÉMICO

El rendimiento académico, por ser multicausal, envuelve una capacidad explicativa de los distintos factores y espacios temporales que intervienen en el proceso de aprendizaje. Existen diferentes aspectos que se asocian al rendimiento académico, entre los que intervienen componentes tanto internos como externos al individuo. Pueden ser de orden social, cognitivo y emocional, que se clasifican en tres categorías: determinantes personales, determinantes sociales y determinantes institucionales, que presentan subcategorías o indicadores (Garbanzo, 2007).

Para vincular de forma directa ambos conceptos utilizados en la investigación podemos señalar lo siguiente a través de una frase simplificada de forma general “Factores que influyen en el rendimiento académico del estudiante universitario “

De acuerdo a otros estudios sustentados en los factores asociados al rendimiento académico, se establece que todos los alumnos llegarán a la universidad con motivación para aprender, pero no es así e incluso si tal fuera el caso, algunos alumnos aún podrían encontrar aburrida o irrelevante la actividad universitaria. Así mismo, el docente en primera instancia debe considerar cómo lograr que los estudiantes participen de

manera activa en el trabajo de la clase, es decir, que generen un estado de motivación para aprender. Por otra parte, pensar en cómo desarrollar en los alumnos la cualidad de estar motivados para aprender, de modo que sean capaces de educarse a lo largo de su vida (Navarro, 2004), y finalmente, que los alumnos participen cognoscitivamente, en otras palabras, que piensen a fondo acerca de qué quieren estudiar (Navarro, 2004). Todo proceso educativo busca mejorar el rendimiento del estudiante, de ahí su importancia y la necesidad de considerar los factores que intervienen en él. Estos factores, también llamados determinantes del rendimiento académico, son difíciles de identificar, sin embargo, requieren acotarse para establecer la influencia e importancia que cada uno tiene en el proceso educativo (Tejedor, 2003).

El rendimiento académico es un término multidimensional determinado por los diversos objetivos y logros pretendidos por la acción educativa. Desde la perspectiva operativa del término, se define como la “nota o calificación media obtenida durante el periodo universitario que cada alumno haya cursado” (Tejedor, 1998). Las variables y los indicadores que inciden en él se clasifican de diversas maneras, a saber: variables demográficas o de identificación (sexo, edad, estado civil, experiencia laboral), variables académicas (tipos de estudios cursados, curso, opción en que se estudia una carrera, rendimiento previo), variables socio-familiares (estudios de los padres, situación laboral de los mismos, lugar de residencia familiar, lugar de estudio).

El presente estudio explica los resultados académicos, a través de la calificación final (promedio ponderado acumulado) y los créditos acumulados en el primer año de estudios en la Universidad Nacional de Piura, en función de las variables no pedagógicas propuestas por Tejedor (2003): Académicas, psicológicas, socio-familiares y de identificación. Por ello, a continuación, se describirán las teorías educativas que analizan la influencia de estas variables sobre el rendimiento académico en una evaluación, en una asignatura, en un conjunto de asignaturas de una institución o de un sistema educativo. Adicionalmente, se examinará la

relación entre las variables pedagógicas y el rendimiento, pues así se tendrá una idea sobre aquellos factores explicativos que se estarían dejando de lado.

Para sustentar las teorías que se presentó, se acudió a los resultados de investigaciones fundamentalmente cuantitativas, aunque también se sintetizó las conclusiones de estudios cualitativos. Se debe señalar que, cuando se trata de identificar y comprender la naturaleza de los factores o variables que son percibidos por estudiantes y/o profesores como potenciales predictores del rendimiento académico, los trabajos cualitativos suelen usar encuestas aplicadas a muestras que no necesariamente poseen representatividad estadística, así como grupos focales y entrevistas a profundidad. En cambio, los estudios cuantitativos acuden a metodologías de tipo experimental y no experimental. A su vez, las investigaciones no experimentales se apoyan en, por lo menos, una de las siguientes tres técnicas de análisis de estadístico: descriptivo, correlacional o regresión.

2.8. INTRODUCCIÓN DE MODELO LOGÍSTICO BINARIO

Dentro del análisis de datos involucran la predicción del valor de resultado de una variable dependiente categórica. Estas incluyen aplicaciones en la medicina para predecir el estado de salud de un paciente, en estudios de mercados la predicción acerca de la aceptación de un producto, en el ambiente pedagógico para predecir el rendimiento académico de un estudiante. La regresión logística binaria es una técnica que puede resultar bastante útil en estas y otras situaciones.

Los modelos de elección discreta son modelos en los que la variable dependiente es de carácter cualitativo (no métrica). Los modelos de elección discreta están muy relacionados con el análisis discriminante. Actualmente existe una tendencia a utilizar en mayor medida los modelos de elección discreta debido a que requieren realizar menos supuestos, lo que permite, en general obtener unos resultados más robustos.

En estadística, la regresión logística es un tipo de análisis de regresión utilizado para predecir el resultado de una variable categórica (una variable que puede adoptar un número limitado de categorías) en función de las variables independientes o predictoras. Es útil para modelar la probabilidad de un evento ocurriendo como función de otros factores. El análisis de regresión logística se enmarca en el conjunto de Modelos Lineales Generalizados (GLM por sus siglas en inglés) que usa como función de enlace la función logit. Las probabilidades que describen el posible resultado de un único ensayo se modelan, como una función de variables explicativas, utilizando una función logística.

La regresión logística es usada extensamente en las ciencias médicas y sociales. Otros nombres para regresión logística usados en varias áreas de aplicación incluyen modelo logístico, modelo logit, y clasificador de máxima entropía. La Regresión Logística es una técnica estadística multivariante que nos permite estimar la relación existente entre una variable dependiente no métrica, en particular dicotómica y un conjunto de variables independientes métricas o no métricas.

La regresión logística está diseñada para emplear una mezcla de variables predictoras categóricas y continuas para predecir una variable categórica de resultado o dependiente. Es muy común verlo como una alternativa al análisis discriminante. Varios de los conceptos que se discuten dentro de este análisis de regresión. A diferencia del análisis discriminante la regresión logística tiene muy pocos y menos estrictos, supuestos, además aun cuando los supuestos del análisis discriminante se han cumplido, la regresión logística se desempeña casi igual de bien.

Los modelos de regresión logística son modelos estadísticos en los que se desea conocer la relación entre: Una variable dependiente cualitativa, dicotómica (regresión logística binaria), en relación con una o más

variables independientes, ya sean cualitativas o cuantitativas, siendo la ecuación inicial del modelo de tipo exponencial, si bien su transformación logarítmica (Logit) permite su uso como una función lineal.

2.8.1. La Ecuación Logística

La regresión logística binaria es una regresión que se aplica con una variable dependiente dicotómica (un atributo cuya ausencia o presencia se ha puntuado con los valores cero y uno, respectivamente), donde la variable dependiente no contiene valores de datos sin procesar, pero en cambio es la oportunidad de que un evento de interés ocurra. La ecuación de regresión logística es:

$$\ln(Odds) = \alpha + B_1 X_1 + \dots + B_K X_K$$

Donde los términos de la derecha son los términos estándar para las variables independientes y la intercepción es una ecuación de regresión. Sin embargo, del lado izquierdo está el logaritmo natural de la oportunidad y la cantidad $\ln(Odds)$ llamada logit. En principio puede variar de menos a más infinito, por lo que se elimina el problema de la predicción fuera de los límites de la variable dependiente. La oportunidad está relacionada con la probabilidad por:

$$Odds = \frac{P_i}{1 - P_i}$$

Existe una relación lineal con las variables independientes en la regresión logística, pero es lineal en el logaritmo natural de la oportunidad y no en las probabilidades originales dado que estamos interesados en la probabilidad de un evento, por ejemplo el código más alto de una variable dicotómica o categoría de interés, la

ecuación logística se puede transformar en otra ecuación en la probabilidad, entonces tiene esta forma:

$$Prob(evento Y_i = 1) = \frac{1}{1 + e^{-(\alpha + \beta_{ki} X_{ki})}}$$

Dónde: Prob (y =1 | X) es la probabilidad de que y tome el valor 1 (presencia de la característica estudiada), en presencia de las covariables X:

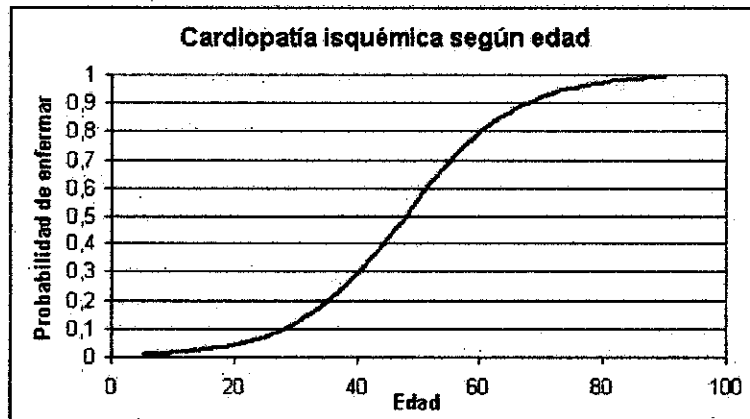
X_{ki} : es un conjunto de k covariables que forman parte del modelo.

α : es la constante del modelo o término independiente.

β_{ki} : Los coeficientes de las covariables.

Esta ecuación no se puede calcular con el método de mínimos cuadrados, en lugar de esto, los parámetros del modelo se estiman con la técnica de máxima verosimilitud. Derivamos los coeficientes que hacen que nuestros valores observados sean los más probables para el grupo de las variables independientes, esto se hace a través de iteraciones internas dentro de los programas estadísticos. La probabilidad que se obtiene es el punto de partida para la clasificación de un elemento cualquiera de la muestra a uno de los grupos definidos dentro de la variable de respuesta.

La ecuación de partida en los modelos de regresión logística es lo que se denomina distribución logística. En la siguiente imagen vemos un ejemplo de esta distribución: la probabilidad de padecer enfermedad coronaria en función de la edad. Como puede verse, la relación entre la variable dependiente (cualitativa dicotómica), y la covariable (edad, cuantitativa continua en este caso), no es definida por una recta (lo que correspondería un modelo lineal), sino que describe una forma sigmoidea (distribución logística).



2.8.2. Elementos del Análisis de Regresión Logística

Cuando se hace un análisis de regresión logística se tiene dos objetivos generales:

1. Alcanzar la más alta precisión predictiva que sea posible con un determinado grupo de variables predictoras seleccionadas a partir de la significancia obtenida por cada una de ellas. También se asume como consecuencia la acción de clasificar individuos dentro de las categorías (presente/ausente) de la variable dependiente, según la probabilidad que tenga de pertenecer a una de ellas dada la presencia de determinadas covariables.
2. Determinar el efecto de un grupo de variables en la probabilidad, además del efecto de las variables tomadas individualmente sobre el resultado global del modelo. Esto significa la ejecución de cuantificar la importancia de la relación existente entre cada una de las covariables y la variable dependiente, lo que lleva implícito también clarificar la existencia de interacción y confusión entre covariables respecto a la variable dependiente (es decir, conocer los Odds ratio para cada covariables).

La clasificación de casos es un asunto simple en la regresión logit, se predice que un caso está en el valor más bajo de la variable dependiente si su probabilidad predicha es menor que 0.5, de otra forma se predice que está en la categoría superior.

2.8.3. Supuestos de la Regresión Logística

La regresión logit, tiene menos supuestos que la regresión estándar, la logística necesita que:

1. Las variables independientes sean intervalares, de razón o dicotómicas.
2. Que se incluya a todas las predictoras relevantes, que no incluya variables irrelevantes y que tengan una relación lineal entre sí.
3. El valor esperado del error es 0.
4. Que no haya autocorrelación.
5. Que no haya correlación entre el error y las variables independientes.
6. Que no haya multicolinealidad entre las variables independientes.

Multicolinealidad: Si bien existen pruebas que permiten comprobar la existencia de colinealidad entre covariables, cabe reseñar aquí que los modelos con multicolinealidad entre las covariables introducidas llamarán la atención por la presencia de grandes errores estándar, y frecuentemente, estimaciones de coeficientes anormalmente elevadas.

Sin embargo la multicolinealidad no afecta al sentido de las estimaciones (la multicolinealidad no hará que aparezca significación donde no la hay, y viceversa).

Se dice que existe multicolinealidad cuando dos o más de las covariables del modelo mantienen una relación lineal. Cuando la colinealidad es perfecta, es decir, cuando una covariable puede determinarse según una ecuación lineal de una o más de las restantes covariables, es posible estimar un único coeficiente de todas las covariables implicadas. En estos casos debe eliminarse la covariable que actúa como dependiente.

Los Estadísticos Odds Ratio

El Estadístico Odds: mide el cociente de probabilidades para una observación i de elegir la opción 1 frente a la opción 0, es decir:

$$Odds = \frac{P_i}{1 - P_i}$$

Este indicador en términos de regresión logística es igual a:

$$Odds = \frac{P_i}{(1 - P_i)} = e^{a + \beta_i X_i}$$

El Estadístico Odds Ratio: Se define como el cociente de Odds de esta manera se tiene:

$$Odds\ Ratio = \frac{\frac{P_i}{(1 - P_i)}}{\frac{P_m}{1 - P_m}}$$

Si lo que se quiere es comparar la utilidad que la opción elegida proporciona al individuo (observación) i , con respecto a la utilidad percibida por el individuo (observación) m , entonces se define como Odds Ratio.

De este modo en términos de regresión logística es:

$$Odds\ Ratio = \frac{\frac{P_{i+1}}{(1 - P_{i+1})}}{\frac{P_i}{(1 - P_i)}} = \frac{e^{\alpha + \beta_i(X_i+1)}}{e^{\alpha + \beta_i X_i}} = e^{\beta_i(X_i+1 - X_i)} = e^{\beta_i}$$

2.8.4. Significatividad estadística de los parámetros estimados

La distribución del estimador del parámetro β es aproximadamente $N(\beta; \sigma_{\hat{\beta}})$. En tal situación, se puede construir un intervalo de confianza del parámetro estimado, para testar si dicho valor es significativamente distinto de cero de forma individual. El contraste a realizar quedaría definido como:

$H_0: \beta = 0$ El parámetro es igual a cero

$H_1: \beta \neq 0$ El parámetro es distinto de cero

El intervalo de confianza proporciona un rango de posibles valores para el parámetro, por lo que si el valor estimado no pertenece a dicho intervalo, se deberá rechazar la hipótesis nula. El intervalo quedaría definido como:

$$\hat{\beta} - Z_{\frac{\alpha}{2}}\sigma_{\hat{\beta}} \leq \beta \leq \hat{\beta} + Z_{\alpha/2}\sigma_{\hat{\beta}}$$

donde α es la probabilidad de que el verdadero valor del parámetro β se halle fuera del intervalo, y z es el valor tabular de la distribución $N(0;1)$ que deja a su derecha una probabilidad igual a $\alpha/2$.

A partir de la expresión anterior se puede fijar un rechazo de la hipótesis nula cuando:

$$\left| \frac{\hat{\beta}}{\sigma_{\hat{\beta}}} \right| \geq Z_{\alpha/2}$$

2.8.5. Medidas de Bondad de Ajuste del modelo

Se sabe que cualquier variable dependiente de otra u otras variables, toma valores según las variables de las que depende. Por otra parte, esa variable dependiente irá tomando valores siguiendo o describiendo una determinada distribución de frecuencias; es decir, toman los valores que tienen las variables independientes, si el experimento se repite múltiples veces, la variable dependiente tomará para un conjunto de variables independientes un determinado valor, y la probabilidad de ocurrencia de dicho valor vendrá dado por una distribución de frecuencias concreta: una distribución normal, una distribución binomial, una distribución Hipergeométrica, etc. En el caso de una variable dependiente dicotómica (como el caso que nos ocupa), la distribución de frecuencias que seguirá será la binomial, que depende de la tasa de éxitos (X sujetos de un total de N , que sería el elemento variable), para un determinado tamaño muestral N y probabilidad $Pr(i)$ de ocurrencia del evento valorado por la variable

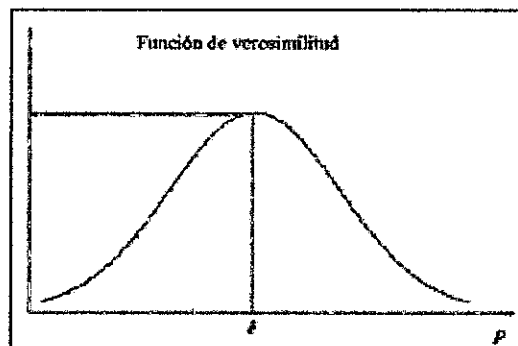
dependiente (parámetros constantes). La función de densidad de esta distribución de frecuencias vendrá dada por la siguiente expresión:

$$\Pr(y) \approx f(x) = \binom{N}{x} p^x (1-p)^{N-x}$$

Si en la expresión anterior introducimos los datos concretos de nuestra muestra de N sujetos (es decir, se convierte el elemento variable X en parámetro), y se hace depender el resultado de la función de densidad del parámetro "probabilidad de ocurrencia" (p , que de esta forma se convierte en variable), se está generando su función de verosimilitud, $f(p | x)$ (función dependiente de p dado el valor muestral de x) o $L(p)$ (L de *likelihood*), que ofrece como resultados las probabilidades de la función de densidad ajustada a los datos:

$$f(p | x) = \binom{N}{x} p^x (1-p)^{N-x}$$

Se deduce que, para una *muestra* concreta, esa probabilidad será diferente según qué valores tome el parámetro "probabilidad de ocurrencia":



Se demuestra que la mejor estimación del parámetro es aquel valor que maximiza esta función de verosimilitud, ya que son estimadores consistentes (conforme crece el tamaño muestral, la estimación se aproxima al parámetro desconocido), suficientes (aprovechan la información de toda la muestra), asintóticamente normales y asintóticamente eficientes (con mínima varianza), si bien no siempre son insesgados (no siempre la media de las estimaciones para diferentes muestras tenderá hacia el parámetro desconocido).

El uso de la función de verosimilitud en la estimación, hace que la bondad del ajuste en los modelos de elección discreta sea un tema controvertido, ya que en estos modelos no existe una interpretación tan intuitiva como en el modelo de regresión clásico. A continuación se describen los contrastes más utilizados en la literatura para medir la bondad de ajuste en un modelo Logit y que concretaremos en: Estadístico de Wald, índice de cociente de verosimilitudes, el estadístico chi-cuadrado de Pearson, el porcentaje de aciertos estimados en el modelo, y la prueba de Hosmer-Lemeshow.

Estadístico de Wald

Este Estadístico de Prueba evalúa la significancia de los coeficientes, se define como el vector matriz de los coeficientes estimados del siguiente modo según las Hipótesis y donde se busca contrastar la proposición de que un coeficiente aislado es distinto de 0, y sigue una distribución normal de media 0 y varianza 1. Su valor para un coeficiente concreto viene dado por el cociente entre el valor del coeficiente y su correspondiente error estándar. La obtención de significación indica que dicho coeficiente es diferente de 0 y merece la pena su conservación en el modelo. En modelos con errores estándar grandes, el estadístico de Wald puede proporcionar falsas ausencias de significación (es decir, se incrementa el error tipo II).

Tampoco es recomendable su uso si se están empleando variables de diseño.

$$H_0: \beta_i = 0, \forall i$$

$$H_1: \text{Al menos un } \beta_i \neq 0$$

Estadístico de Prueba:

$$W = \hat{\beta}' [\hat{I}(\hat{\beta})]^{-1} \hat{\beta} = \hat{\beta}' (X' V X) \hat{\beta} \sim \chi^2_{\alpha, k+1}$$

Donde X y V son las matrices:

$$X = \begin{bmatrix} 1 & x_{11} & \dots & x_{1k} \\ 1 & x_{21} & \dots & x_{2k} \\ \dots & \dots & \dots & \dots \\ 1 & x_{n1} & \dots & x_{nk} \end{bmatrix}_{n \times (k+1)}$$

$$V = \begin{bmatrix} \hat{p}(x_1)(1-\hat{p}(x_1)) & 0 & \dots & 0 \\ 0 & \hat{p}(x_2)(1-\hat{p}(x_2)) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \hat{p}(x_n)(1-\hat{p}(x_n)) \end{bmatrix}_{n \times n}$$

Decisión: si $W > X_{\alpha,k}^2$ rechazamos H_0 con un nivel de significancia fijado α , concluimos que la variable independiente influye en la probabilidad del suceso.

Prueba de Hosmer – Lemeshow

Otra medida global de la exactitud predictiva, no basada en el valor de la función de verosimilitud sino en la predicción real de la variable dependiente, es el contraste de clasificación diseñado por Hosmer y Lemeshow (1989). Dicho contraste consiste en realizar comparaciones entre el valor estimado y el observado por grupos. Para ello las observaciones se dividen en J grupos (generalmente 10) aproximadamente iguales, dividiendo el recorrido de la probabilidad en deciles de riesgo (esto es probabilidad de ocurrencia del fenómeno < 0.1 , < 0.2 , y así hasta < 1). Cada uno de los grupos contiene n_j observaciones, y en cada uno de los J grupos se define:

□ Y_j Como la suma de valores 1 en cada uno de los grupos

$$\frac{P_{11}}{P_{11} + P_{12}}$$

□ \bar{P}_j como la media de los valores predichos en cada grupo

$$\frac{P_{21}}{P_{11} + P_{21}}$$

A partir de esta información se puede construir una tabla de contingencia a través de la que se compara tanto la distribución de ocurrencia, como la de no ocurrencia prevista por la ecuación y los valores realmente observados. El contraste se realiza comparándolas frecuencias observadas y esperadas a través del cálculo del estadístico:

$$HL = \sum_{j=1}^J \frac{(Y_j - n_j \bar{P}_j)^2}{n_j \bar{P}_j (1 - \bar{P}_j)}$$

Reg. 6473 — 15/12/15 LND

Hosmer y Lemeshow demuestran que cuando el modelo es correcto el estadístico HL sigue una distribución chi-cuadrado con J-2 grados de libertad, por lo que valores inferiores del estadístico calculado respecto al teórico indicarán un buen ajuste del modelo.

El uso correcto de este contraste requiere un tamaño de muestra adecuado para asegurar que cada grupo cuenta al menos con cinco observaciones. Además el estadístico chi-cuadrado es sensible al tamaño muestral, permitiendo que esta medida encuentre diferencias estadísticamente muy pequeñas cuando el tamaño muestral crece.

Índice de Cociente de Verosimilitudes:

La función de verosimilitud puede también utilizarse para obtener un estadístico, que tiene cierta semejanza con el coeficiente de determinación calculado en la estimación lineal, conocido “índice de cociente de verosimilitudes”. Este estadístico compara el valor de la función de verosimilitud de dos modelos; uno corresponde al modelo estimado que incluye todas las variables independientes (modelo completo) y el otro sería el del modelo cuya única variable independiente es la constante (modelo restringido). El estadístico, también conocido como R^2 de McFadden ya que fue propuesto por McFadden en 1974, se define como:

$$RV = ICV = 1 - \frac{\log L}{\log L(0)}$$

Donde L es el valor de la función de verosimilitud del modelo completo (el estimado con todas las variables independientes) y L(0) es el valor correspondiente del modelo restringido (el que incluye únicamente en la estimación el término constante).

Se trata de ir contrastando cada modelo que surge de eliminar de forma aislada cada una de las covariables frente al modelo completo. En este caso cada estadístico R.V. sigue una χ^2 con un grado de libertad (no se asume normalidad). La ausencia de significación implica que el modelo sin la covariable no empeora respecto al modelo completo (es decir, da igual su presencia o su ausencia), por lo que según la estrategia de obtención del modelo más reducido (principio de parsimonia), dicha covariable debe ser eliminada del modelo ya que no aporta nada al mismo. Esta prueba no asume ninguna distribución concreta, por lo que es la más recomendada para estudiar la significación de los coeficientes.

El ratio calculado tendrá valores comprendidos entre 0 y 1 de forma que:

Valores próximos a 0 se obtendrán cuando $L(0)$ sea muy parecido a L , situación en la que nos encontraremos cuando las variables incluidas en el modelo sean poco significativas, es decir, la estimación de los parámetros β no mejora el error que se comete si dichos parámetros se igualaran a 0. Por lo que en este caso la capacidad explicativa del modelo será muy reducida.

Cuanto mayor sea la capacidad explicativa del modelo, mayor será el valor de L sobre el valor de $L(0)$, y más se aproximará el ratio de verosimilitud calculado al valor 1.

El estadístico χ^2 de Pearson:

Para medir la bondad del ajuste también se utilizan medidas del error que cuantifican la diferencia entre el valor observado y el estimado. En concreto, para contrastar la hipótesis nula de que:

$$H_0: Y_i = \hat{P}_i ; \text{ lo que equivale a } H_0: Y_i - \hat{P}_i = e_i = 0$$

Se construye un estadístico que recoge los residuos estandarizados o de Pearson del modelo Logit, que se definen como la diferencia entre el valor observado de la variable respuesta y el estimado, dividido por la estimación de la desviación típica, ya que la esperanza es nula. A través del contraste de multiplicadores de Lagrange, se puede calcular el estadístico conocido con el nombre de χ^2 de Pearson, que se define como:

$$\chi^2 = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n \frac{(Y_i - \hat{P}_i)^2}{\hat{P}_i(1 - \hat{P}_i)}$$

Este estadístico es similar a la suma de cuadrados de los residuos del modelo de regresión convencional. El ajuste del modelo será mejor cuanto más cerca esté el valor del estadístico de cero. Para saber a partir de qué valor puede considerarse el ajuste como aceptable es necesario conocer la distribución del estadístico. Éste estadístico, bajo la hipótesis nula, se distribuye como una chi-cuadrado con (n-k) grados de libertad, por lo que su valor se compara con el valor teórico de las tablas de la chi-cuadrado para contrastar la hipótesis nula. Si el valor calculado es superior al valor teórico se rechaza la hipótesis nula lo que equivale a decir que el error cometido es significativamente distinto de cero, es decir, se trataría de un mal ajuste.

Porcentaje de aciertos estimados del modelo:

Otra de las vías utilizadas para determinar la bondad de un modelo Logit es predecir con el modelo los valores de la variable dependiente Y_i de tal manera que $Y_i = 1$ si $\hat{P}_i > c$ ó $Y_i = 0$ si $\hat{P}_i < c$.

Generalmente, el valor que se asigna a c para determinar si el valor de la predicción es igual a 1 o a 0 es de 0.5, puesto que parece lógico que la predicción sea 1 cuando el modelo dice que es más probable obtener un 1 que un 0.

Sin embargo, la elección de un umbral igual a 0.5 no siempre es la mejor alternativa. En el caso en que la muestra presente desequilibrios entre el número de unos y el de ceros la elección de un umbral igual a 0.5 podría conducir a no predecir ningún uno o ningún cero. Así, supuesta una muestra de 1000 observaciones donde 100 son 1 y el resto 0, si el modelo incluye término constante, la media de las probabilidades estimadas en la muestra será 0.1, se deduce que la media de las probabilidades estimadas por el modelo, ha de coincidir con la proporción de unos que haya en la muestra, por lo que será casi imposible que se obtenga un valor estimado superior a 0.5. Si el umbral seleccionado es de 0.5, con esta regla nunca se llegarían a estimar valores iguales a 1. El modo de resolver este problema es tomar un umbral más pequeño. Ejemplo extraído de modelos de elección discreta (Medina, Eva).

Con cualquier tipo de regla predictiva similar se cometerán dos errores: habrá ceros que se clasifiquen incorrectamente como unos y unos que se clasifiquen incorrectamente como ceros. Si se reduce el umbral por debajo de 0.5 aumentará el número de veces que se clasifican correctamente observaciones para las que $Y_i = 1$, pero también aumentará el número de veces en que se clasifiquen observaciones como unos para las que $Y_i = 0$. Cambiando el valor del umbral se reducirá siempre la probabilidad de un error de un tipo y se aumentará la probabilidad del otro tipo de error. Por lo que el valor que debe tomar el umbral depende de la distribución de datos en la muestra y de la importancia relativa de cada tipo de error. Una vez

seleccionado el nivel del umbral, y dado que los valores reales de Y_i son conocidos, basta con contabilizar el porcentaje de aciertos para decir si la bondad del ajuste es elevada o no. A partir de este recuento se puede construir el siguiente cuadro de clasificación:

Donde P_{11} y P_{22} corresponderán a predicciones correctas (valores 0 bien predichos en el primer caso y valores 1 bien predichos en el segundo caso), mientras que P_{12} y P_{21} corresponderán a predicciones erróneas (valores 1 mal predichos en el primer caso y valores 0 mal predichos en el segundo caso). A partir de estos valores se pueden definir los índices necesarios que se utilizan en la clasificación de aciertos del modelo.

Cuadro de clasificación de aciertos

| | | Valor real de Y_i | |
|---------------------------|-----------------|---------------------|-----------|
| | | $Y_i = 0$ | $Y_i = 1$ |
| Predicción de \hat{P}_i | $\hat{P}_i < c$ | P_{11} | P_{12} |
| | $\hat{P}_i > c$ | P_{21} | P_{22} |

Análisis de los Indicadores de la Varianza Explicada

La regresión logística también proporciona dos valores que son análogos, al coeficiente de determinación R cuadrado, de la regresión estándar. Dada la relación para una variable dicotómica, la cantidad de varianza explicada del modelo se debe definir diferente. La R cuadrado de Cox y Snell y la R cuadrado de Nagelkerke, por lo general se prefiere la segunda porque puede llegar a alcanzar un valor máximo de 1. Por cualquiera de estos dos valores su definición es

acerca de lo que las variables independientes pueden alcanzar a explicar de la varianza total del modelo.

Su identificación está relacionada con la prueba F de significancia global del modelo, y esta se origina después de analizar el ajuste a través de las pruebas ómnibus específicas del modelo logit y los resultados obtenidos en ambas pruebas influyen directamente sobre los estudios futuros de la predicción.

Precisión en la Predicción

Un indicador de que tan bien se desempeña el modelo está en su habilidad para clasificar a los casos con precisión en las dos categorías que están definidas en la variable predicha, la precisión predictiva global y las precisiones específicas se obtienen en el cuadro de porcentajes a través de una división del número de casos estimados de forma correcta sobre el total de casos clasificados para ambas categorías conocidos antes de iniciar el trabajo de regresión logit. Estos valores son importantes para generar respuestas inmediatas en cuanto a la buena predicción que se obtiene con el modelo, y debe ser comparada de inmediato con las pruebas de significancia para verificar si ambos objetivos son compatibles o las diferencias son notorias, bajo el concepto inicial de la falta de correspondencia entre el ajuste del modelo y los estadísticos de verosimilitud, o la significancia de las variables individuales y la habilidad predictiva del mismo. Dado que encontrar un modelo significativo no es razón suficiente de tener una elevada predictibilidad.

A los valores de porcentajes particulares para la tabla de aciertos se les conoce como especificidad que se hace en función de la categoría

más común o específica y para la categoría menos común se le conoce como sensibilidad, por ser la categoría sensible o de interés sobre el estudio y utilización del modelo logit.

Realizando predicciones

Con los coeficientes de regresión podemos hacer predicciones sobre los valores de casos individuales, entonces vamos a calcular la probabilidad de la categoría de interés en el estudio, remplazando los valores apropiados individuales en cada variable para dicho caso individual en la función matemática del modelo ya conocido, este resultado se puede expresar en términos de oportunidad con la consideración exponencial que de elevar dicha probabilidad al valor (e) como se discutió para los valores de los coeficientes del modelo. Y la probabilidad por si sola sea utilizada para llevar dicho caso individual a una de las dos categorías de la variable dependiente.

Curva Operativa de Rendimiento (ROC)

Con la tabla de clasificación de aciertos, se describió su significado de los porcentajes encontrados para las dos categorías de la variable predicha, entonces todo parte del conocimiento que el punto de corte en el análisis logit es 0.5, porque así lo define por defecto el programa estadístico y la naturaleza de las investigaciones lo determinan de forma general de esa manera. En la parte ultima de la regresión logit se obtiene la curva ROC, (por sus siglas en inglés) cuyos ejes son la sensibilidad o susceptibilidad (eje y) y el complemento de la especificidad (eje x) y al formar las coordenadas para la curva se obtiene los valores de referencia para el patrón de desempeño de las proporciones, y ver la relación que guarda con la variable de estado.

El uso inicial de esta curva es ver los balances entre sus dos ejes, es probable que se esté interesado en cambiar el punto de corte debido a la experiencia de los investigadores y la investigación empieza nuevamente con una percepción diferente, pero hay que tener en cuenta que esto genera un análisis previo de la distribución inicial de las categorías de estudio.

III. METODOLOGÍA

3.1. TIPO DE INVESTIGACIÓN

3.1.1. Según la naturaleza del objeto de estudio

Empírica: Las ciencias naturales y las ciencias sociales tienen como objeto de estudio los hechos materiales, los fenómenos que son visibles en la realidad; por eso en estas ciencias se realiza investigación empírica, es decir, investigación referida a los hechos observables en la realidad.

3.1.2. Según el método de estudio de las variables

Investigación Cuantitativa: La investigación cuantitativa se realiza cuando el investigador mide las variables y expresa los resultados de la medición en valores numéricos.

3.1.3. Según el número de variables

Multivariada: Las investigaciones multivariadas consideran que el efecto es producido por la concurrencia de dos o más variables independiente que actúan sobre la variable dependiente.

3.1.4. Según el tipo de datos que producen

Primaria: Porque es una fuente directa donde se consiguieron los datos que son la información involucrada en la investigación; esto se refiere a que recopilamos la información a través de un instrumento de medición utilizado como parte de la metodología del trabajo. Que en este estudio de investigación fue La Encuesta.

3.1.5. Según el tiempo de aplicación de la variable

Transversales: El tipo de investigación, de acuerdo a la finalidad es aplicada y de acuerdo a la técnica de contrastación, es correlacional explicativa. El diseño del presente trabajo de investigación es de corte transversal.

3.2. POBLACIÓN DE ESTUDIO

Alumnos ingresantes a la Universidad Nacional de Piura, pertenecientes a la Facultad de Ciencias, en el primer año de estudios 2013; sabiendo de antemano que dicha facultad cuenta con 5 carreras profesionales que son: Matemática, Ingeniería Electrónica, Estadística, Física, y Ciencias Biológicas, al igual que en todos los años en particular la universidad ofrece distintas formas de ingreso, por lo cual es necesario precisar que se cuenta con una población objetivo, bastante considerable en función de su tamaño total que son los alumnos registrados en la bases de datos de la secretaría académica de esta facultad, y además los cursos en su mayoría,

para las distintas especialidades son similares dados que proviene dado en su enfoque general en los dos primeros ciclos de estudios de estas respectivas especialidades, teniendo en cuenta también que al momento de la recolección de datos algunos alumnos ya se han retirado de las especialidades que ingresaron y por lo tanto sería imposible realizar la encuesta sobre ellos, teniendo estas y otras consideraciones particulares obtendremos el número total de alumnos para toda la facultad y por especialidad y así poder definir el tamaño de muestra a utilizar, el resultado final al verificar los datos es una población de 340 alumnos para toda la facultad y las subpoblaciones para cada especialidad es Biología igual a 66, Física igual a 22, Matemática igual a 50, Estadística igual a 46, y Electrónica y Comunicaciones igual a 83.

3.3. SELECCIÓN DE MUESTRA

En los estudio con regresión logística se utiliza donde la investigación es aplicada de forma general se utiliza la fórmula de Freeman $[10 \text{ o } 30 \cdot (k + 1)]$ diez o treinta veces el número de variables o factores independientes que se van a incluir en el análisis, como tamaño de muestra mínimo y máximo respectivamente donde k = número de variables independientes del modelo. Pero este concepto teórico que se considera como fórmula general, nos señala como requisito que se debe tener un tamaño poblacional de al menos 1000 datos o muestras bastantes grandes. Dado que en el caso particular de esta investigación no se cuenta con un tamaño de muestra así, consideramos utilizar muestreo estadístico como forma también de aplicar los conocimientos básicos y fundamentales que comprende nuestra carrera profesional.

Debido a que el interés de la investigación de este estudio se centra en determinar la probabilidad de alumnos desaprobados dentro de la Facultad de Ciencias de la Universidad Nacional de Piura; y a sabiendas que

existen en dicha facultad cinco especialidades. Se considera la población homogénea y por interesarnos estudiar una característica específica en todas ellas como variable principal de estudio denominada dependiente, que puede tomar una de dos posibles categorías (alumno “Aprobado” o alumno “Desaprobado”), se puede utilizar un muestreo estratificado con afijación proporcional.

El muestreo estratificado es un diseño de muestreo probabilístico en el que dividimos a la población en subgrupos o estratos. La estratificación puede basarse en una amplia variedad de atributos o características de la población como edad, género, nivel socioeconómico, ocupación, etc. Así, consideramos una población heterogénea con N unidades, y en la que la subdividimos en L subpoblaciones denominados estratos lo más homogéneas posibles no solapadas, atendiendo a criterios que puedan ser importantes en el estudio, de tamaños N_1, N_2, \dots, N_L . Obviamente $N_1 + N_2 + \dots + N_L = N$, donde N es el total de individuos de la población. La muestra estratificada de tamaño n se obtiene seleccionando una muestra aleatoria simple de tamaño n_h ($h = 1, 2, \dots, L$) de cada uno de los estratos en que se subdivide la población de forma independiente. De igual modo, $n_1 + n_2 + \dots + n_L = n$, donde n es el tamaño de la muestra que queremos seleccionar.

Sabemos también que podemos utilizar muestreo aleatorio simple y luego realizar la distribución proporcional para cada especialidad, puesto que esta es una fórmula más sencilla donde la realización es fácil de entender e interpretar, decidimos utilizarla para encontrar el tamaño de muestra que vamos a seleccionar. Entonces siguiendo esta descripción obtenemos los cálculos de la muestra aleatoria y su afijación proporcional, esto se da a conocer a continuación con los valores de población y subpoblación ya conocidos y considerando los criterios obligatorios en la fórmula de aplicación y se detalla en los siguientes pasos como sigue:

➤ CALCULO DEL TAMAÑO DE LA MUESTRA

Calculamos el tamaño de la muestra se tienen en cuenta las siguientes especificaciones: Nivel de confianza 95%, Error máximo admisible $\pm 2\%$, Probabilidad de Éxito de 0.5. De esta manera se obtiene:

$$n = \frac{NZ^2P(1-P)}{(N-1)E^2 + Z^2P(1-P)}$$

Donde:

n: Tamaño de muestra

N: Tamaño de la población

Z: Valor obtenido en la distribución normal estándar para el nivel de confianza dado.

E: Error máximo admisible.

De esta manera obtenemos:

$$n = \frac{(340)(1.96)^2(0.5)(1-0.5)}{(340-1)(0.02)^2 + (1.96)^2(0.5)(1-0.5)}$$

$$n = 296$$

Donde:

N: 340

Z: 1.96

E: 0.02

P: 0.5

Q: 0.5

Dado que el valor del tamaño de la muestra resulta ser 296 estudiantes y comparado con nuestro conjunto total de alumnos igual a 340 parece ser que no hay suficientes cambios para realizar una selección a través de

muestreo probabilístico simple, pero debemos entender también que se debe efectuar la corrección de la muestra porque primero debemos encontrar el factor de muestreo y verificar si este es superior a 5%, y lo cual haremos a continuación:

$$f = \frac{n}{N}$$

El resultado de f (fracción de muestreo) será igual a:

$$f = \frac{296}{340} = 0.8706$$

Consideramos el resultado del valor de f y podemos notar que lógicamente es superior al 5%, entonces debemos realizar la corrección del tamaño de la muestra para obtener una selección adecuada como lo describen los requisitos de cálculo en el curso de muestreo estadístico.

$$n = \frac{n}{1 + f}$$

Remplazando los valores en la última ecuación encontraremos el tamaño de la muestra a seleccionar, el cual después será repartido profesional para cada especialidad. (Resultado ya redondeado a su valor entero)

$$n = \frac{296}{1 + 0.8706} = 158$$

Con este tamaño de muestra obtenido hacemos nuestra afijación proporcional:

$$N = N_1 + N_2 + N_3 + N_4 + N_5 = 340$$

$$N_1 = \left(\frac{66}{267}\right) \times 158 = 39$$

$$N_2 = \left(\frac{22}{267}\right) \times 158 = 13$$

$$N_3 = \left(\frac{50}{267}\right) \times 158 = 30$$

$$N_4 = \left(\frac{46}{267}\right) \times 158 = 27$$

$$N_5 = \left(\frac{83}{267}\right) \times 158 = 49$$

Donde:

$$N_1 = \textit{Biología}$$

$$N_2 = \textit{Física}$$

$$N_3 = \textit{Matemática}$$

$$N_4 = \textit{Estadística}$$

$$N_5 = \textit{Ingeniería Electrónica}$$

3.4. METODO Y PROCEDIMIENTO DE LA INFORMACIÓN

En este trabajo se utiliza la Regresión Logística Binaria, que es el procedimiento de análisis ideal para evaluar este tipo de investigación que estamos realizando, dado que las características de la investigación son las adecuadas para realizar esta forma de trabajo; en relación al tipo de variables incluidas y sobre todo los objetivos principales y específicos que estamos desarrollando en este estudio de investigación.

El tamaño de muestra ya está calculado de acuerdo a las formulas precisas que se definen en la teorías de los cursos de estadística.

La información quedo registrada en las encuestas, luego se procedió a introducir en programa IBM SPSS 20 generando una base de datos con todas las variables dependientes e independientes, y desde allí se realizó un análisis exploratorio para verificar los datos erróneos, perdidos y en blanco obtenidos al recolectar la información y ejecutar los pasos necesarios para corregir estos posibles errores.

La encuesta está debidamente validado por las reglas estadísticas que acreditaron la confiabilidad de la encuesta, estas reglas son el alfa de cronbach, por lo tanto podemos asumir que la recepción de la información evaluada en su conjunto con todas las circunstancias antes descritas, serán el inicio para introducir los datos al programa estadístico ideal para la evaluación, y empezar a construir el modelo y los resultados que se irán encontrando a partir del análisis de las tablas que arroja este procedimiento.

El método que se utiliza dentro de la regresión logística es el conocido como método introducir, que realiza la inclusión de todas las variables de forma conjunta en un solo y único paso general, identificando cuales de todas las variables son significativas y que van a ser introducidas dentro del modelo en la construcción de la ecuación matemática de la función logística y descartando aquellas que no son significativas.

Las variables son todas aquellas incluidas dentro de la encuesta de trabajo de la investigación, y están definidas dentro del cuadro de resumen de todas variables, en ella se toma en cuenta variables Dicotómicas, en algunos casos variables con más de dos posibles respuestas, variables numéricas, y van a ser redefinidas según las características de SPSS, para un trabajo más consolidado y de mejor comprensión teórica, y práctica.

3.5. DEFINICIÓN Y OPERACIONALIZACIÓN DE VARIABLES

A partir de la información que se obtuvo para el cohorte de alumnos de la Universidad Nacional de Piura que ingresaron en el 2013, se estima un modelo que permitió identificar los principales determinantes del rendimiento del primer año de estudios de los mencionados alumnos. En este modelo, se considera un grupo importante de los factores académicos que, según la investigación educativa, han mostrado tener mayor influencia sobre el desempeño universitario. De esta forma, se buscó identificar aquellos atributos de mayor importancia relativa sobre el rendimiento académico, en presencia de variables de control tales como las características psicológicas, socio-familiares y de identificación de los estudiantes. Se utilizan dos variables diferentes para medir el rendimiento del primer año de universidad:

3.5.1. Variable Dependiente

La calificación acumulada que mide el promedio ponderado del alumno luego de haber cursado su primer año de estudios es la variable principal a considerar porque con esta calificación obtuvimos nuestra variable dependiente en la investigación. En los registros de la UNP, esta nota se denomina Promedio Ponderado Acumulado (PPA). Para calcularla, se consideró el promedio de las notas obtenidas en los cursos que corresponden a cada ciclo de estudios, en este caso los dos primeros ciclos para los estudiantes objeto de esta investigación que corresponden al I y II del año 2013, para luego promediar, nuevamente, los dos resultados acumulados de ambos periodos de estudio en función del creditaje de cada curso. La naturaleza del cálculo y el sistema de calificación usado permiten que se trabaje la variable como una dependiente continua. Sin embargo, la calificación se encuentra acotada

en un rango que va desde el 0 hasta el 20, siendo 0 la menor nota que podría obtener el alumno y 20, la máxima, de acuerdo con la escala vigesimal utilizada por el sistema educativo peruano.

El número de créditos acumulados por el alumno luego de haber finalizado los primeros dos primeros semestres académicos en la UNP. Esto es, la suma final del total de créditos obtenidos luego de haber cursado y aprobado los cursos de cada ciclo. Sabemos también que esta variable es una dependiente continua, con un solo resultado numérico entero que representa la descripción en el cálculo final de los créditos acumulados por cada alumno. Esta variable está incluida en la encuesta pero no será utilizada, solo nos sirve para verificar cual es la disposición que existe en los estudiantes entre el número de créditos aprobados y la calificación promedio ponderada, más adelante tal vez en investigaciones futuras se podría realizar un modelo de regresión logística binaria utilizando esta variable como dependiente y hacer comparaciones con los resultados que encontremos en este proyecto.

3.5.2. Variables Independientes

El modelo de rendimiento académico incorpora un conjunto de variables explicativas entre las que destacan aquellas referidas al rendimiento del alumno y las características sociales y personales que destacan el reconocimiento de patrones definidos a partir de sus respuestas, además de sus indicadores psicológicos (personalidad, inteligencia) y medios o formas utilizados en sus estilos de aprendizaje, así como otras variables socio-familiares y de identificación. También se consideran preguntas donde los estudiantes otorgan una descripción de su opinión frente al desempeño de sus docentes, el lugar donde realizan sus actividades académicas, y una autovaloración de su disposición al aprendizaje y su rendimiento académico actual.

En términos generales, se propone que los resultados académicos del primer año de estudios universitarios mejoran si, durante el colegio, el estudiante obtuvo un mayor rendimiento escolar y siguió algún ciclo preuniversitario. Además, se considera que la formación que ofrecen los colegios privados, y las academias preuniversitarias, contribuye a elevar el rendimiento académico en la Universidad. El alumno tiene formas o métodos para estudiar lo que obtiene en clase y por eso se ha hecho preguntas con respecto a ciertas características, estilos o formas que ellos utilizan en sus métodos de aprendizaje, para las cuales se consideran categorías o posibilidades frecuentes en aquel tipo de situaciones.

Dentro de las variables socio-familiares, se considera el nivel socioeconómico de los alumnos, medido a través de los niveles de ingresos económicos los que son registrados en la base de datos porque son los mismos alumnos que dan a conocer ese tipo de datos, así como también situación familiar de forma cualitativa. Se cree que a mayores ingresos es un buen indicador de los recursos económicos que pueden ser destinados a la educación del alumno. Si los padres poseen mayor poder adquisitivo, podrán reforzar el proceso de aprendizaje de sus hijos mediante bienes (equipos de cómputo, por ejemplo) o servicios (como los de profesores particulares que nivelen el déficit de conocimientos escolares) que contribuirían a elevar el rendimiento en la Universidad.

Todas variables explicativas o independientes que se han tomado en cuenta suman un total de 32 preguntas incluidas en la encuesta, donde cada una de ellas se describe en el cuadro siguiente y se puede leer para tener el conocimiento preciso de su denotación, simbología, y categorías que le pertenecen y lógicamente también la clasificación al tipo de variable a donde corresponden.

OPERACIONALIZACIÓN DE LAS VARIABLES DEPENDIENTES E INDEPENDIENTES DE LA ENCUESTA

Todas las variables tienen su sustento en la parte teórica que se refiere a los factores implicados como necesarios dentro del estudio, es por eso que la tabla muestra las preguntas formuladas y la forma como serán introducidas en el programa estadístico para la ejecución de los procedimientos y las futuras explicaciones.

VARIABLES

| Variable (Nombre) | Definición | Indicador | Valores |
|--|--|--|---|
| <i>Dependiente</i> | | | |
| Situación Académica | Calificación que mide el rendimiento académico inmediato que es resultado del aprendizaje. | Calificación final (promedio) acumulada ponderada al terminar los dos primeros ciclos de estudios. | Aprobado(1); Desaprobado(0) |
| Promedio ponderado de los alumnos. | Es el puntaje de los alumnos al promediar sus calificaciones de los dos semestres académicos | Es una variable cuantitativa continua con valores desde 0 hasta 20 | Valores que asume de Menores de 11: Desaprobado Mayores de 11. Aprobados |
| Número de créditos aprobados por el alumno | Es la variable que describe el número de créditos totales aprobados por | Es una variable cuantitativa continua donde el número es una respuesta sin rango fijo. | Es un valor numérico que establece el alumno como respuesta y se comprueba con el dato de secretaria académica. |

| Variable (Nombre) | Definición | Indicador | Valores |
|--------------------------|---|---|--|
| | alumno durante los dos semestres. | | |
| Independientes | | | |
| Especialidad | Nombre de especialidad de las cinco que cuenta la facultad de Ciencias de la UNP. | Se categorizan según el código asignado por parte de la universidad a cada escuela profesional | Matemática=1 Estadística=2 Ciencias Biológicas=3 Ingeniería Electrónica=4 Física=5 |
| Edad del Alumno | Evalúa el número de años del estudiante al momento de haber ingresado a la universidad. | Para operacionalizar esta variable, se usa la edad del estudiante al momento de haber ingresado a la universidad. | Mayores a 17 años. (la edad es un valor cuantitativo discreto) |
| Sexo | Esta variable determina si el individuo posee sexo masculino o femenino. | Para operacionalizar esta variable, se usa una variable dicotómica que especifica si el alumno es hombre o mujer. | Masculino(0) Femenino(1) |
| Tipo de residencia | La forma como el estudiante lleva acabo sus actividades en | Se ha tomado en cuenta la formas más usuales de respuestas del alumno | Viajas todos los días (1) Cuenta con pensión en Piura (2) |

| Variable (Nombre) | Definición | Indicador | Valores |
|------------------------------|---|--|--|
| | la UNP. | | Vives con un familiar (3) Tu casa familiar está en Piura (4) |
| Estado Civil | Esta variable determina la condición de individuo con respecto a su estado civil. | Para operacionalizar esta variable, se usa el estado civil del estudiante al momento de haber ingresado a la universidad. | Soltero =1 Casado=2 Separado O divorciado=3 Unión libre=4 Otro=5 |
| Financiamientos de Estudios | Esta variable determina quién financia los estudios del individuo. | Para operacionalizar esta variable, se usa la condición que el estudiante describió en el cuestionario universitario. | Beca de la universidad=1 Trabajo propio=2 Ayuda de padres u otras personas=3 Otros medios=4 |
| Situación Laboral | Esta variable determina la situación laboral del estudiante. | Para operacionalizar esta variable, se usa la situación laboral que describió el estudiante durante el desarrollo de sus ciclos de estudios. | Trabaja=1 No trabaja=0 |
| Hijos del Alumno | Esta variable determina si es estudiante tiene hijos. | Para operacionalizar esta variable, se usa la variable dicotómica. | Sí=1 No=0 |
| Situación Económica Familiar | Es la variable acerca de la afirmación del alumno sobre la condición de | Se utiliza categorías que describen esta situación Familiar según el estudiante. | Buena (1) Estable (2) Deficiente (3) Muy mala (4) |

| Variable (Nombre) | Definición | Indicador | Valores |
|---|---|---|---|
| | su Familia. | | |
| Presupuesto mensual del alumno | Describe cuanto tiene el alumno como presupuesto para sus actividades universitarias durante el mes. | Se categoriza en función de cifras de dinero, descrita por el alumno en un rango necesario. | Menos de 100 (1) Entre 100-300 (2) Más de 300 (3) |
| Tipo y Nombre de colegio en Educación secundaria. | Describe el tipo de colegio donde realizo su secundaria el alumno. | Esta cuenta con dos categorías principales y el nombre es para identificación. | Estatal (1) Particular (2) |
| Ciclo Pre U | Esta variable determina si el estudiante llevo algún ciclo de academia preuniversitaria durante su etapa escolar. | Para operacionalizar esta variable se define una variable dicotómica | 0=No 1=Si |
| Números de Ciclos | Esta variable determina el número de ciclos que el estudiante llevo durante su etapa escolar. | Para operacionalizar esta variable, se usa una escala. | Un ciclo=1 Dos ciclos=2 Más de dos=0 |

| Variable (Nombre) | Definición | Indicador | Valores |
|--|--|---|---|
| Características de tu Personalidad en la Secundaria. | Define como se considera el alumno durante la etapa de educación secundaria | Es una variable Psico-Social y está estructurada según criterios teóricos. | Siempre amable (1) Muy exigente (2) Trabajador en Equipo (3) Buen líder (4) Distraído, Irresponsable (5) |
| Características académicas de tu educación secundaria. | Describe algunas características de percepción acerca de las labores académicas recibidas en secundaria. | Se categoriza según patrones de estructura social en cuantos niveles específicos que regulan la educación básica. | Profesores Eficientes (1) Teoría comprensible (2) Aplicación práctica (3) Tecnología necesaria (4) Análisis de Problemáticas (5) |
| Evaluación de tu desarrollo personal y Social en la universidad. | Determina la autoevaluación del alumno acerca de su desarrollo en la universidad. | La variable esta categorizada según criterios, de teoría de desarrollo personal de universitarios. | Responsable en sus labores académicas (1) Participativo en problemas sociales (2) Exigente en derechos estudiantiles (3) Cumple con deberes éticos y morales (4) Sin ningún interés, solo cumple con sus padres (5) |

| Variable (Nombre) | Definición | Indicador | Valores |
|----------------------------------|---|--|--|
| | | | |
| Rendimiento académico Secundaria | Describe el rendimiento académico obtenido por el estudiante durante el nivel secundario. | Variable categorizada en una escala de valores nominales simples. | Muy bueno (1) Bueno (2) Regular (3) Malo (4) Muy malo (5) |
| Número de Alumnos en tu Salón | Esta variable determina el número de alumnos con los que el estudiante llevó su último año escolar. | Para operacionalizar esta variable, se usa una escala. | Menos de 15 alumnos = 1 15-25 alumnos = 2 25-40 alumnos = 3 Más de 40 alumnos = 4 |
| Horario Extra | Esta variable determina si el estudiante tiene un horario personal para estudiar las asignaturas del I ciclo de su carrera. | Para operacionalizar esta variable se usa una variable dicotómica. | SI=1 NO=0 |
| Lugar de Estudio | Esta variable determina el lugar de estudios del individuo. | Para operacionalizar esta variable se usa una variable dicotómica. | SI=1 NO=0 |

| Variable (Nombre) | Definición | Indicador | Valores |
|-----------------------|--|---|--|
| Uso de Materiales | Esta variable determina si el estudiante cuenta y usa material bibliográfico para sus estudios. | Para operacionalizar esta variable se usa una variable dicotómica. | SI=1 NO=0 |
| Frecuencia de Estudio | Esta variable determina la frecuencia del tiempo que dedica el estudiante para poner en práctica lo impartido en las aulas universitarias. | Para operacionalizar esta variable se usa lo descrito por el estudiante en el cuestionario universitario. | Todos los días=1 De vez en cuando=2 Sólo cuando tengo evaluaciones=3 Nunca estudio=4 |
| Estudio para exámenes | Esta variable determina cómo el estudiante se prepara para un examen o un trabajo encargado. | Para operacionalizar esta variable se usa una variable dicotómica. | Consulta solamente la bibliografía asignada por el profesor=1 Se interesa por consultar otras fuentes=2 |
| Nivel educativo Padre | Esta variable determina el actual nivel educativo padre. | Para operacionalizar esta variable se usa escala. | 1= Primaria 2= Secundaria 3= Técnico/tecnológico 4= Pregrado |

| Variable (Nombre) | Definición | Indicador | Valores |
|-----------------------|--|--|---|
| | | | 5= Postgrado |
| Nivel Educativo Madre | Esta variable determina el actual nivel educativo de la madre. | Para operacionalizar esta variable se usa escala. | 1= Primaria 2= Secundaria 3= Técnico/tecnológico 4= Pregrado 5= Postgrado |
| Intervalo de Sueldo | Evalúa el nivel y composición de los Ingresos familiares. | Para operacionalizar esta variable se usa escala. | 1:<1000, 2:1000-1999; 3:2000-2999; 4:>3000. |
| Opinión Familiar | Esta variable determina la opinión familiar con respecto a estudiar en la universidad. | Para operacionalizar esta variable se establece una escala de opiniones. | 1= Hay que estudiar para ser alguien en la vida. 2= Hay que estudiar una carrera que de plata. 3= Hay que estudiar lo que a uno más le gusta. 4 =Para qué perder tiempo en la universidad. |
| Comentario Común | Esta variable determina el concepto que tiene la familia como estudiante. | Para operacionalizar esta variable se establece una escala de opiniones. | 1= Mi familia cree que soy muy buen estudiante. 2 = Mi familia cree que soy buen estudiante. 3 = Mi familia cree que soy un estudiante regular. 4 = Mi familia cree que soy |

| Variable (Nombre) | Definición | Indicador | Valores |
|---|--|---|---|
| | | | mal estudiante. 5 =Mi familia cree que soy pésimo estudiante. |
| Interés Padres | Esta variable determina el interés de los padres para con el estudiante. | Para operacionalizar esta variable se usa una variable dicotómica. | Si=1 NO=0 |
| Consideración personal de alumno | Determina como asume el estudiante su situación académica después del primer año de estudios | Variable categorizada en una escala nominal con valores que describen la situación. | Excelente (1) Buena (2) Regular (3) Mala (4) Pésima (5) |
| Calificación de la enseñanza recibida por los docentes de la UNP. | Se refiere a la evaluación del alumno hacia sus profesores en su desempeño docente | Variable categorizada de forma nominal, las escalas representan las categorías descritas. | Eficiente (1) Buena (2) Regular (3) Mala (4) Pésima (5) |
| Descripción de la situación actual de la UNP. | Es una pregunta que hace énfasis a la situación | Variable de escala nominal con categorías que ajustan una evaluación de la | Excelente institución (1) Problemas sociales y éticos |

| Variable (Nombre) | Definición | Indicador | Valores |
|-------------------|---|--|--|
| | que percibe el alumno acerca de la situación de la UNP. | situación general para una institución superior. | (2) No ayuda a los estudiantes (3) Regularmente responsable (4) |

3.6. CONSTRUCCIÓN DE LA BASE DE DATOS

La población objeto de estudio corresponde al total de ingresantes a la Universidad Nacional de Piura durante el primer año de estudios 2013. Debido a que se tiene varias carreras, se decidió realizar un estudio muestral en consideración de la selección de una muestra correcta tal como lo indica la teoría. El único criterio de inclusión es el período de ingreso de los estudiantes, dado que ese grupo presentaría información académica, psicológica y socio-económica referida a las variables de interés a través de la encuesta que se ha utilizado como instrumento de recolección de datos, donde ellos han dado a conocer sus respuestas de forma individual y secreta. Para la construcción de la base de datos, se obtuvo información, a partir de esas respuestas y es lo que conforman nuestras bases que utilizaremos en esta técnica estadística.

Asimismo, se usaron preguntas para recopilar información académica de los colegios de procedencia de los estudiantes y se encuestó a los alumnos acerca de diversos aspectos de su currícula universitaria mediante la misma encuesta que tenía la misión de lograr la capacidad necesaria de información a recolectar planteada desde el objetivo de la investigación donde se trataba de asignar las preguntas idóneas, de tal forma que se pueda obtener información relevante para los distintos factores descritos, y en consecuencia pueda ser incluida como información válida en la base

de datos que se debe formar una vez realizadas todas las encuestas posibles.

El alumno encuestado proporcionará información de identificación y académica de sí mismo, que se genera cuando el estudiante postula a la UNP, así como la referida a la modalidad de admisión por la cual ingresó y además datos referidos a sus panoramas sociales, familiares, económicos de acuerdo a sus niveles de apreciación cultural, y de los contextos singulares de cada participante.

3.7. PROCEDIMIENTO Y METODO

Para obtener datos de información sobre la variable dependiente que se empleó en la técnica de la encuesta y la investigación es siempre a través de la información documental puesto que esto nos ayuda a corroborar las respuestas de los alumnos con referencia a las dos preguntas incluidas en la encuesta para esta variable. Para cuantificar la variable dependiente (el nivel de rendimiento académico), esta variable será convertida a una variable dicotómica, con dos valores excluyentes entre sí, dependiendo del estado del alumno en función de la nota ponderada obtenida al final de ambos semestres, y sus categoría respectivas son bastantes específicas donde, (1: Desaprobado; 0: Aprobado); en ambos casos la forma de definición viene dada por la calificación numérica obtenida donde (Aprobado \geq 10.5, Desaprobado $<$ 10.5) y es así de forma sencilla como se dio inicio a la formulación de la variable dependiente, la cual fue ingresada en el programa estadístico, además se construyó un modelo en simultaneo en función del número de créditos aprobados, para comparar los resultados obtenidos.

Por otro lado, para cuantificar las variables independientes se realizó una encuesta, con cuestionarios relacionados a las variables o factores

relevantes que se consideran en las hipótesis del presente trabajo de investigación. Además, se utilizó la técnica de recolección indirecta de los datos de información, mediante la revisión de documentos estadísticos, textos de la teoría económica y publicaciones relacionadas al tema de estudio.

La preparación de la información comprendió varios aspectos como la recopilación, clasificación, sistematización u ordenamiento, tabulado y presentación de la información referente a cada variable objeto de estudio; las técnicas que se utilizaron en dichos aspectos son manuales y electrónicas; en el segundo caso se utilizó una computadora, en la cual se le asignará carpetas, archivos y ventanas de acuerdo al orden lógico de las variables y que ello pueda facilitar su guardado, modificación y recuperación de los datos de información procesada, para su posterior análisis ya sea individual o en conjunto. El procesamiento de los datos de información recopilada, se utilizó el paquete estadístico SPSS versión 20.0 y Microsoft Excel.

Finalmente, se indica que la determinación e interpretación de los coeficientes de los factores determinantes (variables independientes) del rendimiento académico (variable dependiente), se utilizó el modelo estadístico, regresión logística binaria o también llamado Logit, pues es método necesario en cuanto a modelos regresión que se ajusta a las cualidades de la información y lo que se pretende dentro de la información, y a partir de los objetivos principales de nuestra investigación, lo cual permitió estimar los coeficientes de cada variable independiente para su respectivo análisis, con el cual se alcanzó las respuestas y las conclusiones adecuadas justificadas en la metodología y sustento teórico de nuestra fundamentación profesional.

3.8. ANÁLISIS DE DATOS

El análisis ha sido efectuado con ayuda del programa SPSS 20, utilizando el módulo de Regresión Logística Binaria. En este caso el SPSS está implementado con tres opciones principales para el análisis los cuales son:

- El método “Introducir”. El cual permite al investigador tomar el mando, decidir que variables se introducen o extraen del modelo.
- El método “Adelante”. Es uno de los métodos automáticos (o por pasos), que deja que el programa vaya introduciendo variables en el modelo, empezando por aquellas que tienen coeficientes de regresión más grandes, estadísticamente significativos. En cada paso reevalúa los coeficientes y su significación, pudiendo eliminar del modelo aquellos que no considera estadísticamente significativos.
- El método “Atrás”. Al igual que el anterior es otro de los métodos automáticos. En este caso parte de un modelo con todas las covariables que se hayan seleccionado en el cuadro de diálogo, y va eliminando del modelo aquellas sin significación estadística.

IV. ANÁLISIS DE LOS RESULTADOS

4.1. CONSTRUCCIÓN DEL MODELO LOGIT

4.1.1. Especificación del Modelo de Regresión Logística

$$P = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_{32} X_{32})}}$$

Hipótesis:

$$H_0: \beta_1 = \beta_2 = \beta_3 = \dots = \beta_{32} = 0$$

$$H_1: \text{Al menos un } \beta_i \neq 0, \forall i = 1, 2, 3, \dots, 32$$

H_0 : Los factores sociales, familiares, económicos y psicológicos ($X_1, X_2, X_3, \dots, X_{32}$) no influyen significativamente en el rendimiento académico de los estudiantes ingresantes a la Universidad Nacional de Piura en la facultad de ciencias durante su primer año de estudios 2013 (Y).

H_1 : Existen al menos algunos factores sociales, familiares, económicos y psicológicos ($X_1, X_2, X_3, \dots, X_{32}$) que si influyen significativamente en el rendimiento académico de los estudiantes ingresantes a la Universidad Nacional de Piura en la facultad de ciencias durante su primer año de estudios 2013 (Y).

4.1.2. Ajuste del Modelo de Regresión Logística

El ajuste en este modelo ha sido efectuado mediante el software estadístico SPSS 20.0; se resumen los resultados a través de los siguientes cuadros que arroja el programa estadístico. Se ha utilizado el método introducir porque es aquel que proporciona mejores resultados en comparación del método por pasos.

Cuadro 4.1: Resumen del procesamiento de los casos

| Casos no ponderados | | N | Porcentaje |
|------------------------|--------------------------|-----|------------|
| Casos seleccionados | Incluidos en el análisis | 158 | 100,0 |
| | Casos perdidos | 0 | ,0 |
| | Total | 158 | 100,0 |
| Casos no seleccionados | | 0 | ,0 |
| Total | | 158 | 100,0 |

El cuadro de resumen del procesamiento de los casos nos da a conocer el total de la muestra seleccionada y procesada; obteniendo que ningún dato haya sido excluido. Por lo tanto se debe proseguir con el análisis de resultados, al tener conocimiento que no hay datos perdidos, y no habrá problemas con respecto a ese supuesto necesario para proseguir con el análisis de la regresión Logit.

Cuadro 4.2: Codificación de la variable dependiente

| Valor original | Valor interno |
|----------------|---------------|
| APROBADO | 0 |
| DESAPROBADO | 1 |

La codificación que por defecto efectúa el software estadístico SPSS 20.0 no ha tenido que realizar cambios internos dado que en la matriz

de datos la variable dependiente se ha encontrado codificada como:
Aprobado (0), Desaprobado (1).

Bloque 0: Bloque Inicial

El programa IBM SPSS 20.0 nos proporciona algunos resultados básicos bajo el título de “Bloque 0: Bloque Inicial”. Aunque este modelo no es interesante, nos muestra alguna información básica.

Cuadro 4.3: Tabla de clasificación

| | | | Pronosticado | | |
|--------|--------------------|-----------|--|-------------|------------------------|
| | | | SITUACION REAL DEL ALUMNO SEGUN PROMEDIO PONDERADO | | Porcentaje correcto |
| | | | APROBADO | DESAPROBADO | |
| Paso 0 | SITUACION REAL DEL | APROBADO | 110 | 0 | 100,0 |
| | ALUMNO SEGUN | DESAPROBA | 48 | 0 | ,0 |
| | PROMEDIO | DO | | | |
| | PONDERADO | | | | |
| | Porcentaje global | | | | 69,6 |

Esta tabla de clasificación (Cuadro 4.3), Notamos que se indica la cantidad de alumnos de forma original en las dos categorías de la variable dependiente que vienen a ser 110 Aprobados y 48 Desaprobados lo cual nos indica de forma global en base al resultado porcentual que en el modelo siempre se predice la categoría más común (Aprobado: Alumnos aprobados en su primer año de estudios en la UNP) y esto es correcto para el 69.6% de la muestra. Esto nos predice correctamente a todos los alumnos que han aprobado su primer año de estudios en la UNP, pero sin tomar en cuenta a todos los alumnos que han desaprobado. Esto genera una línea base para lo que se va a evaluar en modelos posteriores. Este resultado porcentual es un valor clave porque nos proporciona la idea puntual de cuanto

estamos dispuestos aceptar en la capacidad predictiva del modelo a través de la tabla de porcentajes correctos, un valor superior a ese resultado generar una mejora de la clasificación con la utilización de la técnica estadística.

Las tablas adicionales en este bloque inicial presentan información sobre las variables en la ecuación (sólo la constante) y las que no están en ella. Estas se presentan en el Anexo de esta investigación. Estos cuadros no son de carácter importante es por eso que solo se describen en la parte ultima del trabajo para referencia en algún momento.

Bloque 1: Introducir

La probabilidad de los resultados observados, dados los cálculos del parámetro, se conoce como verosimilitud. Es una función matemática que maximiza los coeficientes de la ecuación de regresión logística y que por consecuencia minimiza los errores, una lectura parecida al análisis de regresión estándar, la forma de desarrollo es a través de algoritmos iterativos por computadora, Por lo general se utiliza -2 veces el logaritmo natural de la verosimilitud (-2LL) como una medida del ajuste del modelo, dado que tiene vínculos con la distribución de chi-cuadrado. Un buen modelo que tiene una elevada verosimilitud se traduce en un valor pequeño de -2LL. En un ajuste perfecto, -2LL sería igual a 0, y este resultado se comprueba a través de la significancia de la prueba, cuando se ingresa más variables al análisis entonces los cambios que suceden en la verosimilitud final del modelo haría cada vez más cercana la diferencia a cero y eso generaría mejores condiciones en el ajuste del modelo en función de los coeficientes de la ecuación.

**Cuadro 4.4: Pruebas ómnibus sobre los coeficientes
del modelo**

| | | Chi-cuadrado | gl | Sig. |
|--------|--------|---------------------|-----------|-------------|
| Paso 1 | Paso | 67,875 | 46 | ,020 |
| | Bloque | 67,875 | 46 | ,020 |
| | Modelo | 67,875 | 46 | ,020 |

La chi-cuadrado del Modelo es una prueba estadística de la hipótesis nula de que los coeficientes para todos los términos en el modelo son 0, esto es igual a la suposición de la tabla ANOVA de la regresión lineal donde se realizaba la prueba bajo la hipótesis de que los coeficientes del modelo son cero en contraste con que alguno de ellos es diferente, lo cual es la hipótesis formulada al inicio de esta justificación de resultados. El valor encontrado es 67.875 del Cuadro 4.4, es la diferencia entre -2LL inicial (no se muestra en este cuadro pero lo podemos ubicar en el cuadro del anexo de la investigación, pero dicho valor encontrado se basa en el modelo que contiene solo a la constante) y el -2 LL final que contiene a todos los posibles coeficientes del modelo y que también puede ser observado en el siguiente cuadro que se presenta. Tiene 46 grados de libertad, que representa la diferencia entre el número de parámetros en los dos modelos. Por lo tanto rechazamos la hipótesis nula porque la significancia es baja, 0.02 (< 0.05); y concluimos que el grupo de variables que se introducen al modelo de regresión logística binaria para esta investigación mejora la predicción del logaritmo natural de las oportunidades.

Cuadro4.5: Resumen del modelo

| Paso | -2Log de la verosimilitud | R-cuadrado de Cox y Snell | R-cuadrado de Nagelkerke |
|-------------|--------------------------------------|--------------------------------------|-------------------------------------|
| 1 | 126,164 | ,349 | ,494 |

Con todas las variables en el modelo, la bondad de ajuste del estadístico -2LL final es 126.164, que se muestra en el cuadro 4.5 y el cual fue utilizado en el cálculo de la prueba anterior sobre prueba ómnibus de los coeficientes del modelo. Este estadístico de ajuste no suele interpretarse directamente porque no se incluye en la interpretación puntual de una prueba estadística relacionada con la eficacia del modelo pero si es útil para encontrar los resultados de diferencia en la prueba anterior.

La regresión logística también nos muestra dos valores que son análogos a la R^2 en la regresión estándar. Dada la relación entre la media y la desviación estándar para una variable dicotómica, la cantidad de varianza explicada por el modelo se debe definir diferente. La R^2 de Cox y Snell es igual 0.349 y la R^2 de Nagelkerke es igual a 0.494. Por lo general se prefiere esta última porque puede llegar a tomar un valor máximo de 1. A través de cualquiera de estos valores, las variables independientes solo pueden explicar una cantidad mínima de la varianza, si observamos los valores porcentuales de esos coeficientes de determinación no superan el 50%, por eso se concluye que el conjunto de variables elegido no es suficiente para poder establecer la dispersión de los resultados, cuando esto sucede es necesario saber que existen otros factores o variables de acuerdo a las características de la investigación que no han sido considerados en el análisis que podrían determinar de manera más precisa la variabilidad del modelo.

Cuadro 4.6: Prueba de Hosmer y Lemeshow

| Paso | Chi cuadrado | gl | Sig. |
|------|--------------|----|------|
| 1 | 6,760 | 8 | ,563 |

Cuadro 4.7: Tabla de contingencias para la prueba de Hosmer y Lemeshow

| | | SITUACION REAL DEL ALUMNO SEGUN PROMEDIO PONDERADO - APROBADO | | SITUACION REAL DEL ALUMNO SEGUN PROMEDIO PONDERADO - DESAPROBADO | | Total |
|--------|----|---|----------|--|----------|-------|
| | | Observada | Esperada | Observada | Esperada | |
| Paso 1 | 1 | 16 | 15,988 | 0 | ,012 | 16 |
| | 2 | 16 | 15,606 | 0 | ,394 | 16 |
| | 3 | 13 | 14,973 | 3 | 1,027 | 16 |
| | 4 | 15 | 14,084 | 1 | 1,916 | 16 |
| | 5 | 13 | 12,865 | 3 | 3,135 | 16 |
| | 6 | 12 | 11,744 | 4 | 4,256 | 16 |
| | 7 | 11 | 10,322 | 5 | 5,678 | 16 |
| | 8 | 9 | 7,713 | 7 | 8,287 | 16 |
| | 9 | 5 | 5,684 | 11 | 10,316 | 16 |
| | 10 | 0 | 1,019 | 14 | 12,981 | 14 |

Los resultados de la prueba de bondad de ajuste de Hosmer y Lemeshow se muestran en el cuadro 4.6. La prueba estadística se calcula dividiendo los casos en diez grupos de tamaño aproximadamente igual a partir de las probabilidades calculadas, y luego comparando el número de valores observados con los esperados o predichos, en cada categoría de la variable dependiente. La bondad de ajuste es 6.760, y se distribuye como un valor de chi-cuadrado con una significancia de 0.563. Al comparar los eventos observados con los esperados en el contexto de evaluar la bondad de ajuste, esperábamos que dicha significancia sea superior al 5 %, y como la prueba es no significativa dado que se rechaza la hipótesis nula, entonces podemos indicar que los eventos esperados y observados están cerca; lo que implica que el modelo tiene un buen ajusten (> 0.05 ; buen ajuste). En este caso el modelo tiene un buen ajuste, lo que confirma el cambio en la prueba -2LL (prueba del modelo). Observamos en la tabla de contingencias que la división se ha realizado para los diez grupos que se desea en la prueba aunque debemos ser cautelosos con la conclusión dado que las frecuencias para la categoría de interés son inferiores a cinco en varios de los

grupos, esto hace suponer que se necesita una muestra más grande para poder generalizar esta propiedad de ajuste del modelo en base a las probabilidades calculadas.

Cuadro 4.8: Tabla de las variables incluidas dentro de la ecuación del modelo

| | B | S.E. | Wald | df | Sig. | Exp(B) | I.C. 95% para EXP(B) | |
|--|--------|-----------|-------|----|-------|-----------|-------------------------|----------|
| | | | | | | | inferior | superior |
| Paso 1 | | | | | | | | |
| Nombre de la especialidad | -,749 | ,292 | 6,582 | 1 | ,010 | ,473 | ,267 | ,838 |
| Situación económica familiar según el alumno | -1,646 | ,626 | 6,924 | 1 | ,009 | ,193 | ,057 | ,657 |
| Cantidad de presupuesto mensual del alumno | -1,004 | ,477 | 4,428 | 1 | ,035 | ,366 | ,144 | ,933 |
| Opinión familiar sobre su carrera | 1,009 | ,446 | 5,121 | 1 | ,024 | 2,742 | 1,145 | 6,569 |
| Constante | 10,172 | 41621,990 | ,000 | 1 | 1,000 | 26163,053 | | |

El modelo original está en términos del logaritmo natural de las oportunidades o logit. Por lo tanto, el coeficiente B es el efecto de una unidad de cambio en una variable independiente sobre el logaritmo natural de las oportunidades. Interpretando la variable: situación económica familiar según el alumno, el efecto es disminuir el logaritmo natural de las oportunidades en 1.646, ya que la variable está codificada como 1 ó 0. Por otro lado par8ba llegar a interpretar en términos de probabilidad; analizamos la columna Exp(B) presenta el valor exponencial de B. Para la variable: situación económica familiar según el alumno, este valor es de 0. 193, que equivale a

$e^{-1.646}$. Este ahora se expresa en términos de la razón de oportunidad, de manera que si la familia tiene una excelente situación económica, estimamos que la oportunidad de que el alumno salga desaprobado en su primer año de estudios disminuye en un factor de 0.193.

4.2. MODELO FINAL

De acuerdo a los resultados, el análisis con regresión logística en SPSS mediante el método “Introducir” con el criterio de se ha obtenido el siguiente modelo:

Ecuación:

$$P = \frac{1}{1 + e^{10.1172 - 0.749X_5 - 1.646X_{10} - 1.004X_{11} + 1.009X_{29}}}$$

Dónde:

X_5 : *Nombre de la especialidad.*

X_{10} : *Situación económica familiar según el alumno.*

X_{11} : *Cantidad de presupuesto mensual del alumno.*

X_{29} : *Opinión familiar sobre su carrera.*

Los coeficientes β_i son estadísticamente significativos al nivel del 5%. El modelo se ajusta correctamente de acuerdo a la prueba de Bondad de Ajuste de Hosmer y Lemeshow, realizada anteriormente. Esta ecuación se utiliza para el cálculo de probabilidades y es necesaria para encontrar el porcentaje de aciertos del modelo que se realiza a continuación.

4.3. ESTIMACIÓN DEL PORCENTAJE DE PERSONAS ENCUESTADAS CLASIFICADAS CORRECTAMENTE CON EL MODELO LOGIT.

Cuadro 4.9: Tabla de clasificación

| Observado | Pronosticado | | | |
|---|--|-------------|------------------------|------|
| | SITUACION REAL DEL ALUMNO SEGUN PROMEDIO PONDERADO | | Porcentaje correcto | |
| | APROBADO | DESAPROBADO | | |
| Paso 1 SITUACION REAL DEL ALUMNO SEGUN PROMEDIO PONDERADO | APROBADO | 95 | 15 | 86,4 |
| | DESAPROBADO | 15 | 33 | 68,8 |
| Porcentaje global | | | | 81,0 |

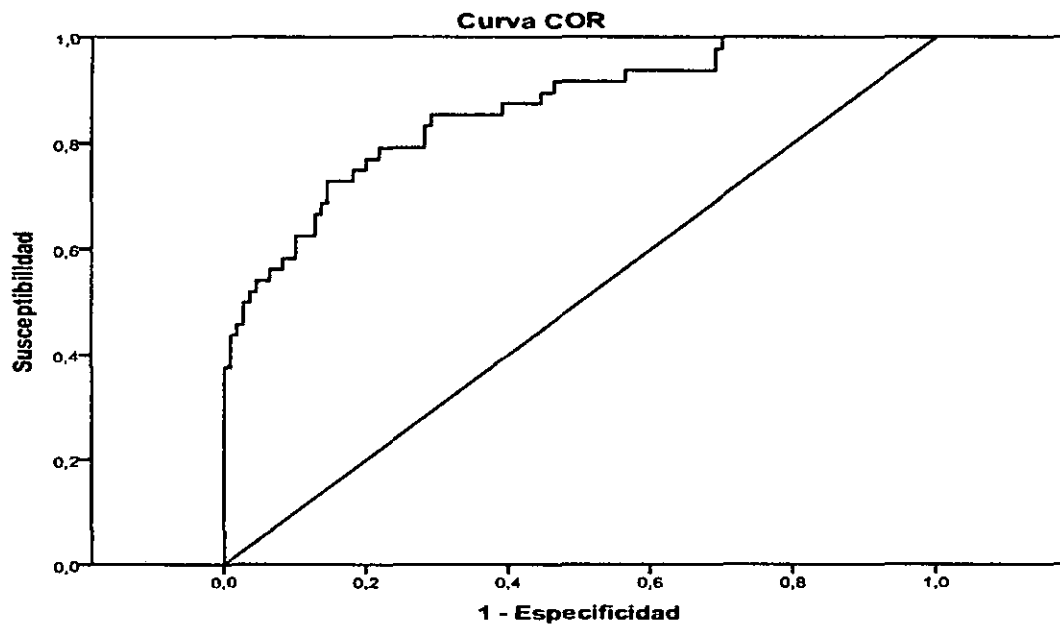
La precisión predictiva global es de 81.0% en el modelo. También observamos que el número de alumnos que salen aprobados en su primer año de estudios son pronosticados correctamente en un 86.4% a este resultado se le conoce como especificidad o susceptibilidad por ser la categoría más común; y sobre el número de alumnos que salen desaprobados en su primer año de estudios son pronosticados correctamente en un 68.8% a este otro valor se le conoce como la sensibilidad porque está dado sobre la categoría de interés o categoría sensible de la variable dependiente. También es de tener en cuenta que el porcentaje global es altamente influenciado siempre por la sensibilidad del modelo. Entonces cuando se consigue una sensibilidad muy alta la capacidad predictiva del modelo será mucho mejor. Con el modelo Logit se encuentra la sensibilidad del mismo que necesitamos y se llama así dado que viene sujeto a la categoría de interés dentro de la investigación. Este valor se calcula de forma simple al igual que para la categoría Especificidad; en otras palabras es el resultado de los alumnos desaprobados o aprobados clasificados correctamente por el modelo sobre

el número de alumnos totales en la condición original de las categorías de la variable dependiente sin haber empleado la herramienta estadística. Entonces esos resultados se pueden encontrar con la siguiente operación de división con los valores de las tablas ya conocidos y se muestran a continuación para conocer el valor porcentual:

Sensibilidad: $33/48=68.8\%$

Especificidad: $95/110= 86.4\%$

4.4. Estimación del valor de corte óptimo: Curva COR (Curva Operativa de Rendimiento)



Área bajo la curva

Variables resultado de contraste:

Probabilidad pronosticada

| Área |
|------|
| ,862 |

El eje vertical registra la sensibilidad (proporción de verdaderos positivos) y el eje horizontal representa la proporción de los falsos positivos (1-

especificidad), es así como se le conoce de forma teórica a los dos ejes que se presentan en la curva ROC y esta gráfica sirve para visualizar como es el comportamiento de la función logística para este conjunto de datos y factores considerados. La línea diagonal proporciona una referencia del patrón para saber si ambos ejes aumentan en la misma proporción, cuanto más se aleja la curva de la línea diagonal se dice entonces que los resultados de predicción son mejores.

En el cuadro posterior a esta gráfica se aprecia que el área bajo la curva es 0.862 la cual difiere en mucho de 0.5 que sería el mínimo exigible para un método de diagnóstico. La diferencia que se presenta es en relación al punto de corte elegido y definido para la clasificación en ambas categorías de la variable dependiente. De acuerdo a los resultados obtenidos se puede afirmar que el área bajo la Curva COR es significativamente mayor que lo mínimo exigible 0.5.

Además en el cuadro siguiente se puede observar que el punto de corte 0.3703310 es el resultado de capacidad predictiva que proporciona un resultado distinto, algo mejor en función de la categoría de interés y la categoría específica de esta variable, dado que se obtiene una Sensibilidad de 72.9% igual a la ya encontrada, pero sucede una disminución muy pequeña en la Especificidad con un valor de 85.5%. Esta situación lo da a conocer la tabla posterior, llamada coordenadas de la curva con el valor sombreado que produce dichos cambios. Esto hace suponer que se logra mejoras en la predicción al elegir un valor diferente de 0.5, la elección de ese punto se hace al observar valores inferiores y superiores a dicho número y al verificar los resultados porcentuales obtenidos según la tabla, se decide por aquel que produce un cambio de mejora en el resultado de sensibilidad o especificidad respectivamente.

Cuadro 4.10: Coordenadas de la curva

Variables resultado de contraste: Probabilidad
pronosticada

| Positivo si es mayor o igual que | Sensibilidad | 1 - Especificidad |
|--|--------------|----------------------|
| ,0000000 | 1,000 | 1,000 |
| ,0000000 | 1,000 | ,991 |
| ,0000000 | 1,000 | ,982 |
| ,0000000 | 1,000 | ,973 |
| ,0000000 | 1,000 | ,964 |
| ,0000000 | 1,000 | ,955 |
| ,0000000 | 1,000 | ,945 |
| ,0000000 | 1,000 | ,936 |
| ,0000000 | 1,000 | ,927 |
| ,0000000 | 1,000 | ,918 |
| ,0000000 | 1,000 | ,909 |
| ,0000000 | 1,000 | ,900 |
| ,0002280 | 1,000 | ,891 |
| ,0015528 | 1,000 | ,882 |
| ,0028188 | 1,000 | ,873 |
| ,0050298 | 1,000 | ,864 |
| ,0089832 | 1,000 | ,855 |
| ,0111835 | 1,000 | ,845 |
| ,0123561 | 1,000 | ,836 |
| ,0146258 | 1,000 | ,827 |
| ,0162259 | 1,000 | ,818 |
| ,0168109 | 1,000 | ,809 |
| ,0172633 | 1,000 | ,800 |
| ,0181311 | 1,000 | ,791 |
| ,0196737 | 1,000 | ,782 |
| ,0225102 | 1,000 | ,773 |
| ,0272971 | 1,000 | ,764 |
| ,0329400 | 1,000 | ,755 |
| ,0365816 | 1,000 | ,745 |
| ,0382727 | 1,000 | ,736 |
| ,0400080 | 1,000 | ,727 |
| ,0422750 | 1,000 | ,718 |
| ,0449924 | 1,000 | ,709 |
| ,0472587 | 1,000 | ,700 |
| ,0515096 | ,979 | ,700 |
| ,0552541 | ,979 | ,691 |
| ,0558436 | ,958 | ,691 |
| ,0566054 | ,938 | ,691 |

Cuadro 4.10: Coordenadas de la curva

Variables resultado de contraste: Probabilidad pronosticada

| Positivo si es mayor o igual que | Sensibilidad | 1 - Especificidad |
|----------------------------------|--------------|-------------------|
| ,0578325 | ,938 | ,682 |
| ,0587007 | ,938 | ,673 |
| ,0605226 | ,938 | ,664 |
| ,0631248 | ,938 | ,655 |
| ,0645722 | ,938 | ,645 |
| ,0676846 | ,938 | ,636 |
| ,0734471 | ,938 | ,627 |
| ,0773327 | ,938 | ,618 |
| ,0786809 | ,938 | ,609 |
| ,0806091 | ,938 | ,600 |
| ,0840945 | ,938 | ,591 |
| ,0889583 | ,938 | ,582 |
| ,0940241 | ,938 | ,573 |
| ,0985416 | ,938 | ,564 |
| ,1014065 | ,917 | ,564 |
| ,1039239 | ,917 | ,555 |
| ,1063300 | ,917 | ,545 |
| ,1086303 | ,917 | ,536 |
| ,1175669 | ,917 | ,527 |
| ,1257836 | ,917 | ,518 |
| ,1277792 | ,917 | ,509 |
| ,1365204 | ,917 | ,500 |
| ,1461053 | ,917 | ,491 |
| ,1479357 | ,917 | ,482 |
| ,1514782 | ,917 | ,473 |
| ,1576371 | ,917 | ,464 |
| ,1631292 | ,896 | ,464 |
| ,1680562 | ,896 | ,455 |
| ,1705774 | ,896 | ,445 |
| ,1730809 | ,875 | ,445 |
| ,1801786 | ,875 | ,436 |
| ,1855646 | ,875 | ,427 |
| ,1867207 | ,875 | ,418 |
| ,1891751 | ,875 | ,409 |
| ,1947862 | ,875 | ,400 |
| ,1992020 | ,875 | ,391 |
| ,2021838 | ,854 | ,391 |
| ,2091038 | ,854 | ,382 |

Cuadro 4.10: Coordenadas de la curva

Variables resultado de contraste: Probabilidad
pronosticada

| Positivo si es mayor o igual que | Sensibilidad | 1 - Especificidad |
|--|--------------|----------------------|
| ,2140828 | ,854 | ,373 |
| ,2189666 | ,854 | ,364 |
| ,2246715 | ,854 | ,355 |
| ,2296192 | ,854 | ,345 |
| ,2359023 | ,854 | ,336 |
| ,2426210 | ,854 | ,327 |
| ,2468680 | ,854 | ,318 |
| ,2478013 | ,854 | ,309 |
| ,2497091 | ,854 | ,300 |
| ,2508959 | ,854 | ,291 |
| ,2546554 | ,833 | ,291 |
| ,2607825 | ,833 | ,282 |
| ,2663588 | ,813 | ,282 |
| ,2700706 | ,792 | ,282 |
| ,2720784 | ,792 | ,273 |
| ,2761582 | ,792 | ,264 |
| ,2795524 | ,792 | ,255 |
| ,2888888 | ,792 | ,245 |
| ,2992346 | ,792 | ,236 |
| ,3032780 | ,792 | ,227 |
| ,3057374 | ,792 | ,218 |
| ,3062098 | ,771 | ,218 |
| ,3161912 | ,771 | ,209 |
| ,3334980 | ,771 | ,200 |
| ,3417231 | ,750 | ,200 |
| ,3425348 | ,750 | ,191 |
| ,3428329 | ,750 | ,182 |
| ,3457105 | ,729 | ,182 |
| ,3498407 | ,729 | ,173 |
| ,3548880 | ,729 | ,164 |
| ,3641004 | ,729 | ,155 |
| ,3703310 | ,729 | ,145 |
| ,3715481 | ,708 | ,145 |
| ,3841440 | ,688 | ,145 |
| ,4016145 | ,688 | ,136 |
| ,4142182 | ,667 | ,136 |
| ,4297641 | ,667 | ,127 |
| ,4412131 | ,646 | ,127 |

Cuadro 4.10: Coordenadas de la curva

Variables resultado de contraste: Probabilidad pronosticada

| Positivo si es mayor o igual que | Sensibilidad | 1 - Especificidad |
|----------------------------------|--------------|-------------------|
| ,4545134 | ,625 | ,127 |
| ,4677881 | ,625 | ,118 |
| ,4767435 | ,625 | ,109 |
| ,4898137 | ,625 | ,100 |
| ,5040084 | ,604 | ,100 |
| ,5118944 | ,583 | ,100 |
| ,5141829 | ,583 | ,091 |
| ,5162761 | ,583 | ,082 |
| ,5338661 | ,563 | ,082 |
| ,5509834 | ,563 | ,073 |
| ,5534515 | ,563 | ,064 |
| ,5558795 | ,542 | ,064 |
| ,5637764 | ,542 | ,055 |
| ,5727132 | ,542 | ,045 |
| ,5751990 | ,521 | ,045 |
| ,5765852 | ,521 | ,036 |
| ,5829802 | ,500 | ,036 |
| ,5905208 | ,500 | ,027 |
| ,5945253 | ,479 | ,027 |
| ,5983278 | ,458 | ,027 |
| ,6003238 | ,458 | ,018 |
| ,6028352 | ,438 | ,018 |
| ,6143174 | ,438 | ,009 |
| ,6315534 | ,417 | ,009 |
| ,6592689 | ,396 | ,009 |
| ,6804776 | ,375 | ,009 |
| ,7000312 | ,375 | ,000 |
| ,7189435 | ,354 | ,000 |
| ,7372756 | ,333 | ,000 |
| ,7562031 | ,313 | ,000 |
| ,7588826 | ,292 | ,000 |
| ,7675994 | ,271 | ,000 |
| ,7891144 | ,250 | ,000 |
| ,8247903 | ,229 | ,000 |
| ,8755805 | ,208 | ,000 |
| ,9213584 | ,188 | ,000 |
| ,9583434 | ,167 | ,000 |
| ,9816120 | ,146 | ,000 |

Cuadro 4.10: Coordenadas de la curva

Variables resultado de contraste: Probabilidad pronosticada

| Positivo si es mayor o igual que | Sensibilidad | 1 - Especificidad |
|----------------------------------|--------------|-------------------|
| ,9919522 | ,125 | ,000 |
| 1,0000000 | ,104 | ,000 |
| 1,0000000 | ,083 | ,000 |
| 1,0000000 | ,021 | ,000 |
| 1,0000000 | ,000 | ,000 |

4.5. Discusión Final

Los resultados finales después de haber realizado el procedimiento de regresión logística binaria a través del programa estadístico SPSS versión 19; nos muestra resultados pronosticados con un buen modelo de predicción. Al ingresar 32 factores que representan al tipo de variables implicadas a los estudiantes ingresantes a la Universidad Nacional de Piura en el año 2013; que en definitiva afectan su realización de sus actividades universitarias en el contexto de evaluar la implicancia que tiene estas sobre el rendimiento académico, las preguntas hechas tratan de responder características principales de identificación, sociales, culturales, económicas acerca de su persona y su ambiente que lo rodea en cuanto a su desarrollo como profesional, se espera que del total de factores ingresados la mayoría de ellas resulten significativas, esto es que tengan influencia directa sobre la variable de respuesta y aquella conclusión se debe dar en base a la respuesta de los indicadores estadísticos, sin embargo se puede notar que solo 4 factores son las que resultan importantes para considerar después de este análisis conjunto ya que estas variables resultaron significativas para el modelo procesado en este proyecto de investigación.

El presente trabajo se obtuvo como finalidad esclarecer si los factores que determinan el rendimiento académico de los estudiantes universitarios es el grado de satisfacción percibida por ellos mismos y que puede ser identificada mientras se verifica las notas conseguidas en el año de evaluación, o diversas variables como lo son: el sexo, el semestre el que se encuentra cursando y el programa académico en el que están inscritos. Si se alinea el contexto general de este tipo de investigaciones con la preparación de esta tesis, es necesario describir que la causa inmediata de llevar a cabo su aplicación corresponde al querer saber cómo afecta los diferentes conceptos sociales, familiares, económicos, en los estudiantes de la UNP. Se dice que el rendimiento académico no es el producto de una única capacidad, sino el resultado sintético de una serie de factores que actúan en, y desde, la persona que aprende.

Debemos considerar que los valores numéricos muestran que riesgo de oportunidad es un factor que genera influencia en la probabilidad de encontrar alumnos desaprobados, entonces la discusión final debe ser que el estudio de los factores agrupados debe tener una agrupación más selectiva, los factores deben ser separados según el tipo que cada uno de ellos representa porque de esa forma se realiza la mayoría de trabajos y realizar modelos múltiples que generen variables independientes en cada situación estudiada por factor.

V. CONCLUSIONES Y RECOMENDACIONES

5.1. CONCLUSIONES

1. Los factores y variables que estadísticamente resultan significativas para el cálculo de la Probabilidad (probabilidad de que el estudiante ingresante a la facultad de ciencia de la Universidad Nacional de Piura en el año 2013;salga desaprobado) mediante el modelo de regresión logística binaria son:

- a) Nombre de la especialidad.
- b) Situación económica familiar según el alumno.
- c) Cantidad de presupuesto mensual del alumno.
- d) Opinión familiar sobre su carrera.

2. De acuerdo a los resultados, el análisis con regresión logística en SPSS mediante el método “Introducir” con el criterio de se ha obtenido el siguiente modelo:

Ecuación:

$$P = \frac{1}{1 + e^{-10.1172 - 0.749X_5 - 1.646X_{10} - 1.004X_{11} + 1.009X_{29}}}$$

Dónde:

X_5 : *Nombre de la especialidad.*

X_{10} : *Situación económica familiar según el alumno.*

X_{11} : *Cantidad de presupuesto mensual del alumno.*

X_{29} : *Opinión familiar sobre su carrera.*

Los coeficientes β_i son estadísticamente significativos al nivel del 5%. El modelo se ajusta correctamente de acuerdo a la prueba de Bondad de Ajuste de Hosmer y Lemeshow, realizada anteriormente. Los valores numéricos representan el aumento o disminución de la probabilidad individual calculada para cada elemento evaluado, y su expresión numérica esta dado en la ecuación matemática del modelo.

La precisión predictiva global es de 81.0% en el modelo. También observamos que el número de alumnos que salen aprobados en su primer año de estudios son pronosticados correctamente en un 86.4% a este resultado se le conoce como especificidad o susceptibilidad por ser la categoría más común; y sobre el número de alumnos que salen desaprobados en su primer año de estudios son pronosticados correctamente en un 68.8% a este otro valor se le conoce como la sensibilidad porque está dado sobre la categoría de interés o categoría sensible de la variable dependiente.

3. Por otro lado para llegar a interpretar en términos de la razón de oportunidad como un valor de riesgo o protección para la probabilidad de encontrar alumnos desaprobados; en primer lugar situación económica familiar según el alumno, este valor es de 0.193, que equivale a encontrar el resultado de la siguiente operación matemática $e^{-1.646}$; este ahora se expresa en términos de la razón de oportunidad, de manera que si la familia tiene una excelente situación económica, estimamos que la oportunidad de que el alumno salga desaprobado en su primer año de estudios disminuye en 5.18 veces en comparación de las otras categorías. De la misma manera los factores llamados Cantidad de Presupuesto mensual del Alumno y Nombre de la Especialidad también tienen valores de odds ratio inferiores a la unidad dado que sus coeficientes son negativos y estos se interpretan de la igual forma, los valores obtenidos son 0.366 y 0.473 respectivamente, entonces también son razones de protección sobre la variable dependiente y disminuyen la oportunidad de encontrar alumnos desaprobados en 2.73 y 2.14 veces. El único factor donde su coeficiente fue positivo es Opinión Familiar sobre su carrera

donde el resultado numérico es igual a 2.742 este es un factor donde la oportunidad funciona como una razón de riesgo sobre la variable dependiente y aumenta la posibilidad de encontrar alumnos desaprobados en un valor de 2.7 veces.

- 4.El punto de corte 0.3703310 es el resultado de capacidad predictiva que proporciona un resultado distinto, algo mejor en función de la categoría de interés y la categoría específica de esta variable, dado que se obtiene una Sensibilidad de 72.9%, pero sucede una disminución muy pequeña en la Especificidad con un valor de 85.5%. Esta situación lo da a conocer la tabla llamada coordenadas de la curva en la justificación de resultados con el valor sombreado que produce dichos cambios. Esto hace suponer que se logra mejoras en la predicción al elegir un valor diferente de 0.4 y se debe considerar su evaluación en futuras investigaciones.

5.2. RECOMENDACIONES

- 1.Las categorías de los factores independientes o variables explicativas como se expresa en la teoría de la regresión logística binaria, no tiene restricciones puesto que ellas pueden definir la inclusión de variables cualitativas o cuantitativas, pero lo más recomendable es obtener factores donde la disposición de la categorías sean de acuerdo a la variable dependiente, y de esa forma lograr mejores interpretaciones, en esta investigación las preguntas tiene múltiples tipos de respuestas; se debe estudiar minuciosamente estos factores con ayuda de profesionales en estadística que generen propuestas para la utilización de nuevas categorías y realizar el procedimiento de la técnica de regresión, para posiblemente lograr resultados diferentes y poder comparar con los resultados ya conseguidos en esta investigación.
- 2.El procedimiento estadístico en todos sus pasos que lo conforman muestran un efectivo desarrollo con resultados en los indicadores de contraste

eficaces, pero debemos considerar las siguientes situaciones, el tamaño de la muestra debería ser más grande se podría extender para más especialidades de la universidad el mismo tipo de diseño para lograr aquel requisito indispensable en la regresión logística binaria, dado que todas las pruebas de contraste se hacen bajo esa condición. La prueba de bondad de Ajuste y la prueba de coeficientes del modelo, así como también la tabla de clasificación de porcentajes correctos necesitan siempre una significativa disposición en cantidad de elementos para ambas categorías de la variable dependiente, y así entonces estas pruebas se desarrollen correctamente.

3.Existen algunas investigaciones parecidas que se han realizado con la misma técnica estadística multivariante, en las conclusiones de aquellos proyectos se ha obtenido resultados que pueden ser elegidos para establecer un punto de corte distinto al genérico, puesto que asumir que la probabilidad de que un alumno universitario en cualquiera de sus menciones sea desaprobado como 0.5 es un valor no tan lógico y que no se ajuste a la realidad objetiva, entonces si discutimos dichas investigaciones y elegimos por otro punto de corte en virtud de las ideas de especialistas en este tipo de trabajos, los resultados podrían ser mejores, además el criterio y el método de introducir las variables se debe disponer como un módulo por pasos hacia adelante dado que son mejores aunque en esta investigación el método introducir generaba los mejores resultados de clasificación.

VI. BIBLIOGRAFÍA

- García, L. (1989). Factores que inciden en el rendimiento académico de los alumnos de la Universidad Nacional de Educación a Distancia (UNED) española (elaboración de un índice). Revista de Tecnología Educativa. 11 (1), 69-95.
- Cueto, S., Jacoby, E. y Pollitt, E. (1997) Factores predictivos del rendimiento escolar en un grupo de alumnos de escuelas rurales. Educación. 6 (12), 213-229.
- Castellanos, L., González, C., González, A. y Manzano, I. (1998). Las matemáticas empresariales: Estudio de los factores determinantes del rendimiento académico. Asociación Española de Profesores Universitarios de Matemáticas para la Economía y la Empresa. Documento disponible en <http://www.uv.es/asepuma/VI/17.PDF>
- Daugherty, T. y Lane, E. (1999). A longitudinal study of academic and social predictors of college attrition. Social Behavior and Personality. Nueva Zelanda. 4 (27), 355-361.
- Cortez, R. (2001). El atraso escolar en el Perú. Lecciones para una agenda de política pública. Centro de Investigación Universidad del Pacífico. Disponible en www.consortio.org/cies/html/pdfs/pm0052.pdf.
- Brunner, J., y Elacqua, G. (2003). Factores que inciden en una educación efectiva: Evidencia internacional. En: Hevia, R. (Ed.). La Educación en Chile, Hoy. Chile: Ediciones Universidad Diego Portales. (www.educoas.org/portal/bdigital/laeducacion/139/pdfs/139pdf1.pdf)
- Reyes, J. (2003). Relación entre el rendimiento académico, la ansiedad ante los exámenes, los rasgos de personalidad, el autoconcepto y la asertividad

en estudiantes del primer año de psicología de la UNMSM”. Tesis para optar por el título profesional de Psicólogo. Lima: Facultad de Psicología de la Universidad Nacional Mayor de San Marcos.

- Tejedor, F. (2003). El poder explicativo de algunos determinantes del rendimiento en los estudios universitarios. *Revista Española de Pedagogía*. 61 (224), pp. 5-32.
- Caro, D. (2003). ¿Cómo mejorar el desempeño académico de los estudiantes que asisten a escuelas en las zonas pobres del Perú? Vásquez, E. y Winkelried, D. (Ed). *Buscando el Bienestar de los pobres: ¿Cuán lejos estamos?*, 1era edición. Lima: Centro de investigación de la Universidad del Pacífico. pp 176-207.
- Vildoso, V. (2003). Influencia de los hábitos de estudio y la autoestima en el rendimiento académico de los estudiantes de la Escuela Profesional de Agronomía de la Universidad Nacional Jorge Basadre Grohmann.
- Cueto, S. (2004). Factores predictivos del rendimiento escolar, deserción e ingreso a educación secundaria en una muestra de estudiantes de zonas rurales del Perú. *Education Policy Analysis Archives*. 12 (35). (<http://epaa.asu.edu/epaa/v12n35/>).
- Hakkinen, I. (2004). Do university entrance exams predict academic achievement? *Economic Studies*. Finlandia. Uppsala University. (83), 33-67.
- Ministerio de Educación del Perú (2004). Factores asociados al rendimiento estudiantil. Unidad de Medición de la Calidad Educativa del MINEDU. Documento de trabajo N° 9. (<http://www2.minedu.gob.pe/umc/admin/images/publicaciones/doctrab/doc9.pdf>)

- Jiménez, C. (2006). Educación familiar y alumnos con alto rendimiento. *Revista Española de Pedagogía*. 64 (234), 273-300.
- Ministerio de Educación del Perú. (2006). La Universidad en el Perú. Razones para una reforma universitaria. Informe 2006. Disponible en www.minedu.gob.pe/dcu/files/libro7.pdf
- Tejedor, F. y García-Valcárcel, A. (2007). Causas del bajo rendimiento del estudiante universitario (en opinión de los profesores y alumnos). *Propuestas de mejora en el marco del EEES. Revista Educación*. 342, 443-473.
- Torres, M., Lajo, R., Campos, E. y Riveros, M. (2007). Rendimiento académico de los alumnos de una facultad de educación de una universidad pública de Lima y su percepción de la calidad académica de los docentes. *Revista de Investigación en Psicología*. 10 (1), 71-89.
- Valdivia, M. y León, G. (2007) School characteristics and academic achievement in Perú: Is the geographical distribution of resources reinforcing social exclusion? (www.grade.org.pe/noticias-/School%20characteristics.pdf)
- Mejía, E. y Gargurevich, R. (2008). Estudio exploratorio sobre buenas prácticas educativas en los colegios de mayor rendimiento en el Pronóstico de Potencial Universitario (PPU) durante los años 2003-2007. Universidad Peruana de Ciencias Aplicadas. Informe no publicado.
- Pérez, P. (2008) *Psicología educativa*. Tercera Edición. Lima: Editorial San Marcos.

VI. ANEXO DE LA INVESTIGACIÓN

Anexo: CUESTIONARIO

La presente encuesta ha sido diseñada para recabar información que contribuya a mejorar el proceso de enseñanza. La información de este cuestionario será tratada únicamente para los fines de la investigación. Gracias por tu apoyo.

Código del estudiante:

Edad:

1. Lugar de residencia:

2. Situación del desarrollo de actividades universitarias en función de tu residencia

☐ Viajas todos los días

☐ Tienes pensión o habitación en Piura

☐ Vives con un familiar en la ciudad

☐ Tú casa familiar esta en Piura

3. Sexo:

| | |
|---|---|
| M | F |
|---|---|

4. Carrera:

☐ 1. Matemática

☐ 2. Estadística

☐ 3. Ciencias Biológicas

☐ 4. Ingeniería Electrónica

☐ 5. Física

5. Estado civil

☐ 1. Soltero (a)

☐ 2. Casado (a)

☐ 3. Separado (a)/divorciado (a)

☐ 4. Unión libre

☐ 5. Otro

6. ¿Cómo financia sus estudios?

- ☐ 1. Beca de la universidad
- ☐ 2. Trabajo propio
- ☐ 3. Ayuda de padres u otras personas
- ☐ 4. Otros medios

7. ¿Cuál es su situación laboral actual?

- ☐ 1. Trabaja
- ☐ 2. No trabaja

8. ¿Usted tiene hijos?

- ☐ 1. Sí ¿Cuántos? ____
- ☐ 2. No

9. ¿Cómo consideras la situación de tu economía Familiar?

- ☐ Excelente
- ☐ Buena
- ☐ Deficiente
- ☐ Muy mala

10. Indica el presupuesto con el cual cuentas en aproximado, para tus estudios universitarios de forma mensual.

- ☐ Menos de 100
- ☐ Entre 100-300
- ☐ Más de 300

11. Escribe el nombre de tu colegio(en el que concluiste el 5to de secundaria)
Menciona si este es Estatal o Privado:

.....
.....

12. Cursaste algún ciclo en una academia preuniversitaria(Piura, o algún otra ciudad):

- ☐ Si
- ☐ No

¿Cuántos ciclos de academia cursaste?

| | |
|----------|--|
| Un ciclo | |
|----------|--|

| | |
|------------|--|
| Dos ciclos | |
| Más de dos | |

13. ¿Cómo consideras la educación que recibes en la universidad?

☐ Excelente

☐ Buena

☐ Deficiente

☐ Muy mala

14. ¿en qué condiciones recibes la educación universitaria (Aulas, Talleres, Infraestructura, etc.)?

☐ Excelente

☐ Buena

☐ Deficiente

☐ Muy mala

15. ¿Cómo consideras la educación que recibes en la universidad?

☐ Excelente

☐ Buena

☐ Deficiente

☐ Muy mala

16. En promedio, ¿cuántos alumnos estudiaban contigo en tu salón de clases durante tu último año de educación secundaria?

☐ Menos de 15

☐ Entre 15-25

☐ Entre 25-40

☐ Más de 40

17. ¿Cómo fue su rendimiento académico durante secundaria?

☐ 1. Muy Bueno

☐ 2. Bueno

☐ 3. Regular

☐ 4. Malo

☐ 5. Muy Malo

18. ¿Usted tiene establecido un horario de estudios fuera del horario de las clases?
- ☐ Sí ☐ No
19. ¿Usted dispone en su casa de un lugar adecuado y exclusivo para estudiar?
- ☐ Sí ☐ No
20. ¿Usted dispone de una computadora e internet y materiales de estudio para la carrera en su casa?
- ☐ Sí ☐ No
21. ¿Con qué frecuencia usted estudia por iniciativa propia?
- ☐ Todos los días
- ☐ De vez en cuando
- ☐ Sólo cuando tengo evaluaciones
- ☐ Nunca estudio
22. Cuando tiene un trabajo o examen de la universidad, usted:
- ☐ Consulta solamente la bibliografía asignada por el profesor
- ☐ Se interesa por consultar otras fuentes
23. ¿Cuál es el nivel educativo de su padre?
- ☐ 1. Primaria
- ☐ 2. Secundaria
- ☐ 3. Técnico/tecnológico
- ☐ 4. Pregrado
- ☐ 5. Postgrado
24. ¿Cuál es el nivel educativo de su madre?
- ☐ 1. Primaria
- ☐ 2. Secundaria
- ☐ 3. Técnico/tecnológico
- ☐ 4. Pregrado
- ☐ 5. Postgrado
25. ¿Cuál es la profesión e ingreso de su padre y de su madre?
- ☐ 1. Profesión del padre: _____ Ingreso: _____

☐ 2. Profesión de la madre: _____ Ingreso: _____

26. De los siguientes enunciados, ¿podría decirnos con cuáles se identifica su familia?

- ☐ Hay que estudiar para ser alguien en la vida
- ☐ Hay que estudiar una carrera que de plata
- ☐ Hay que estudiar lo que a uno más le gusta
- ☐ Para qué perder tiempo en la universidad

27. ¿Sus padres se interesan para que usted alcance buen desempeño en la universidad?

- ☐ Sí
- ☐ No

28. En términos generales, su familia qué concepto tiene de usted como estudiante.

- ☐ Mi familia cree que soy muy buen estudiante
- ☐ Mi familia cree que soy buen estudiante
- ☐ Mi familia cree que soy un estudiante regular
- ☐ Mi familia cree que soy mal estudiante
- ☐ Mi familia cree que soy pésimo estudiante

29. Como consideras tu situación actual después de tu primer año de estudios

- ☐ Excelente
- ☐ Buena
- ☐ Regular
- ☐ Mala
- ☐ Pésima

30.Cuál es tu promedio ponderado después de los semestres llevados, solo considera el valor máximo entero en ti descripción. (por ej. 12.42= 12).

31. Cuantos créditos aprobados obtuviste al finalizar los dos semestres llevados, en el año anterior, en conjunto de créditos obligatorios y electivos.

32. Cuál es tu apreciación en cuanto a la educación superior recibida por parte de tus docentes, su preparación, y el desarrollo de su trabajo.

☐ Eficiente

☐ Buena

☐ Regular

☐ Mala

☐ Pésima

33. ¿Cómo manifiestas en tu apreciación personal en cuanto a los actos de corrupción por las autoridades, de acuerdo a la gestión que desempeñan?

☐ Excelente

☐ Buena

☐ Deficiente

☐ Muy mala

34. ¿en cuanto a seguridad social y contaminación ambiental, desarrollo personal como calificas los ambientes universitarios?

☐ Excelente

☐ Buena

☐ Deficiente

☐ Muy mala